

SVG in Wikipedia and Wikimedia Commons



Brion Vibber
Wikimedia Foundation

SVG Open - October 4, 2009
Mountain View, CA



Wikimedia Foundation



“ Imagine a world in which every single human being can freely share in the sum of all knowledge. That's our commitment. ”


Broadly speaking, we're here to create educational and reference materials, distributed in open formats under free licenses, in many languages -- and we build them collaboratively over the internet to get the most bang for your buck.

Wikipedia

Welcome to **Wikipedia**,
the **free encyclopedia** that **anyone can edit**.
3,051,603 articles in **English**

[Overview](#) · [Editing](#) · [Questions](#) · [Help](#)

Today's featured article



Murray Chotiner (1909–1974) was an American **political strategist**, **attorney**, government official, and close associate and friend of **President Richard Nixon** during much of Nixon's political career. He served as **campaign manager** for the future president's run for the **United States Senate** in 1950 and for his vice presidential bid in 1952, and managed the campaigns of other **California Republicans**. He was active in each of Nixon's two successful runs for the **White House** in low-profile positions. After Congress investigated Chotiner in 1956, suspecting the attorney was using his connections to Nixon for **influence peddling** to benefit his private clients, the Vice President and his former campaign manager temporarily parted ways. Nixon recalled him to work on his 1962 **gubernatorial** campaign and again for his successful **1968 presidential bid**. After Nixon was inaugurated in 1969, Chotiner received a political appointment to a government position and, in 1970, became a member of the **White House staff**. He returned to private practice a year later, but was involved in **Nixon's 1972 re-election campaign**. He remained an informal adviser to Nixon until he died in **Washington D.C.** following an auto accident in January 1974, and Nixon mourned the loss of a man he described as a counselor and friend. ([more...](#))

Recently featured: [Chinese classifier](#) – [The Slave Community](#) – [Byzantine navy](#)

[Archive](#) – [By email](#) – [More featured articles...](#)




Our flagship site, and of course our most well-known, is Wikipedia, started in 2001. The English-language edition alone is home to 3 million encyclopedia articles with 860 thousand images and other media files.

Wikimedia Commons

Picture of the day

Picture of the day




A flower is on a piece of aerogel which is suspended over a bunsen burner. [+/-](#) (en)

[RSS feed](#) · [By email](#) · [Other Featured pictures](#)

Commons reaches 5 million files

On September 2, the number of files on Wikimedia Commons passed 5,000,000. The five millionth file, a scan of the front page of the Danish newspaper "Kjøbenhavnsposten" from November 28, 1838 (*shown*), was uploaded on 11:46 UTC. Read a [press release by Wikimedia Foundation](#) about the event.



Featured pictures & Quality images

If you are browsing Commons for the first time, you may want to start with [Featured pictures](#), [Quality images](#) or [Valued images](#). You can also see some work created by our highly skilled contributors in [Meet our photographers](#) and [Meet our illustrators](#).

Content

By topic
Nature
[Animalia](#) · [Fossils](#) · [Landscapes](#) · [Marine life](#) · [Plantae](#) · [Weather](#)
Society · Culture
[Art](#) · [Belief](#) · [Coats of arms](#) · [Entertainment](#) · [Events](#) · [Flags](#) · [Food](#) · [History](#) · [Language](#) · [Literature](#) · [Music](#) · [Objects](#) · [People](#) · [Places](#) · [Politics](#) · [Sports](#)
Science

[\[+\] Categories](#)
[\[+\] Commons](#)
[\[+\] Copyright statuses](#)
[\[+\] Media types](#)
[\[+\] Pictures and images by source](#)
[\[+\] Topics](#)
[\[+\] User](#)

In 2004, we launched the Wikimedia Commons project to serve as a central repository for freely-licensed images and media files to be shared directly between all language editions of Wikipedia and other sites. Today, Commons is home to 5 million media files, including over a quarter million SVG images. Though a minority in raw count, SVGs are very frequently used today, making up more than half of the actual images served to our readers.

4 years of SVG

Year 2009...



MSIE 8



Firefox 3.5



Safari 4



Opera 10



Chrome 3

svgweb +
Flash

Native SVG

We added support for uploading SVG graphics to Wikipedia in 2005, released generally in MediaWiki 1.4. SVG is pretty well supported by today's browsers, with native SVG in all major browsers except IE, and ongoing work on adaptors such as the Flash-backed svgweb promising to fill the gap.

4 years of SVG

Year 2005...



MSIE 6



Firefox 1.0



Safari 2



Opera 8

Adobe plug-in

SVG 1.1-Tiny

2005 was not so lucky a time; even the latest Opera had only limited SVG support, and other browsers were dependent on installing the Adobe plugin. With such unreliable client support we started simple, with static images rasterized to flat PNG images on the server. But if we weren't going to serve SVG to clients, why use it in the first place?

Non-destructive editing



http://commons.wikimedia.org/wiki/File:Example_of_an_open_educational_resource_Editing_an_svg_file.ogg

First, Wikipedia is built collaboratively; both articles and images go through multiple revisions and are touched by many authors. A vector format like SVG can be edited non-destructively over multiple generations.

Authoring tools



SVG is a standard format with free and for-pay authoring tools available on multiple platforms. Anyone who wants to contribute can download a program like Inkscape and modify their files -- or create diagrams by hand or programmatically.



WIKIPEDIA
The Free Encyclopedia

Navigation

- [Main Page](#)
- [Community portal](#)
- [Current events](#)
- [Recent changes](#)
- [Random page](#)
- [Help](#)
- [Donations](#)

Search

Toolbox

- [What links here](#)
- [Related changes](#)
- [Special pages](#)

Other languages

- [Deutsch](#)
- [Español](#)

[article](#)

[discussion](#)

[edit this page](#)

[history](#)

Supply and demand

In [microeconomic theory](#), the [partial equilibrium supply and demand economic model](#) originally developed by [Alfred Marshall](#) attempts to describe, explain, and [predict](#) changes in the [price](#) and quantity of [goods](#) sold in [competitive markets](#). The model represents a first approximation for describing a market that is not perfectly competitive. It formalizes the theories used by some economists before Marshall and is one of the most fundamental models of some modern economic schools, widely used as a basic building block in a wide range of more detail and theories. The theory of [demand](#) is important for schools understanding of that it is an explanation of which many [resource allocation](#) made. However, unlike general models, supply schedules and [equilibrium](#) model are forces.



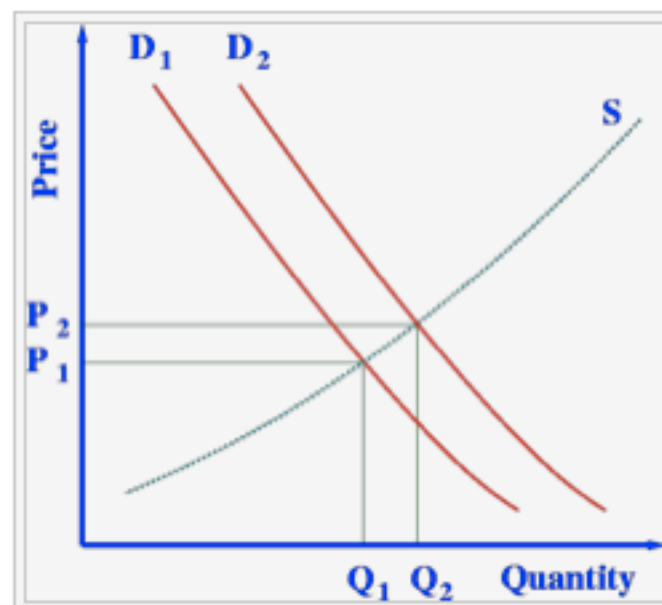
WIKIPEDIA
The Free Encyclopedia

Navigation

- [Main page](#)
- [Contents](#)
- [Featured content](#)
- [Current events](#)
- [Random article](#)

Interaction

- [About Wikipedia](#)
- [Community portal](#)
- [Recent changes](#)
- [Contact Wikipedia](#)
- [Donate to Wikipedia](#)
- [Help](#)



Scaling

[Try Beta](#) [Log in](#) / [create account](#)

Page

[Discussion](#)

Read

[Edit](#)

[View history](#)

Supply and demand

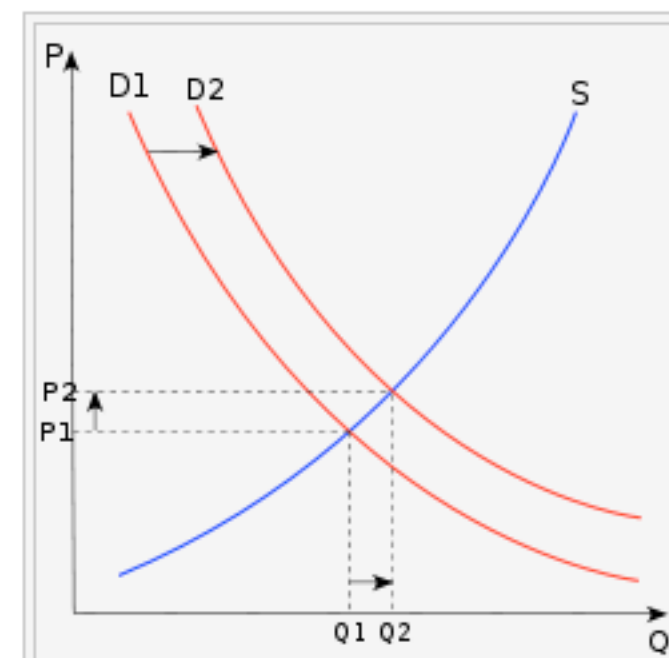
From Wikipedia, the free encyclopedia

For other uses, see [Supply and demand \(disambiguation\)](#).

Supply and demand is an [economic model](#) based on price, utility and quantity in a [market](#). It concludes that in a [competitive market](#), price will function to equalize the quantity demanded by consumers, and the quantity supplied by producers, resulting in an [economic equilibrium](#) of price and quantity.

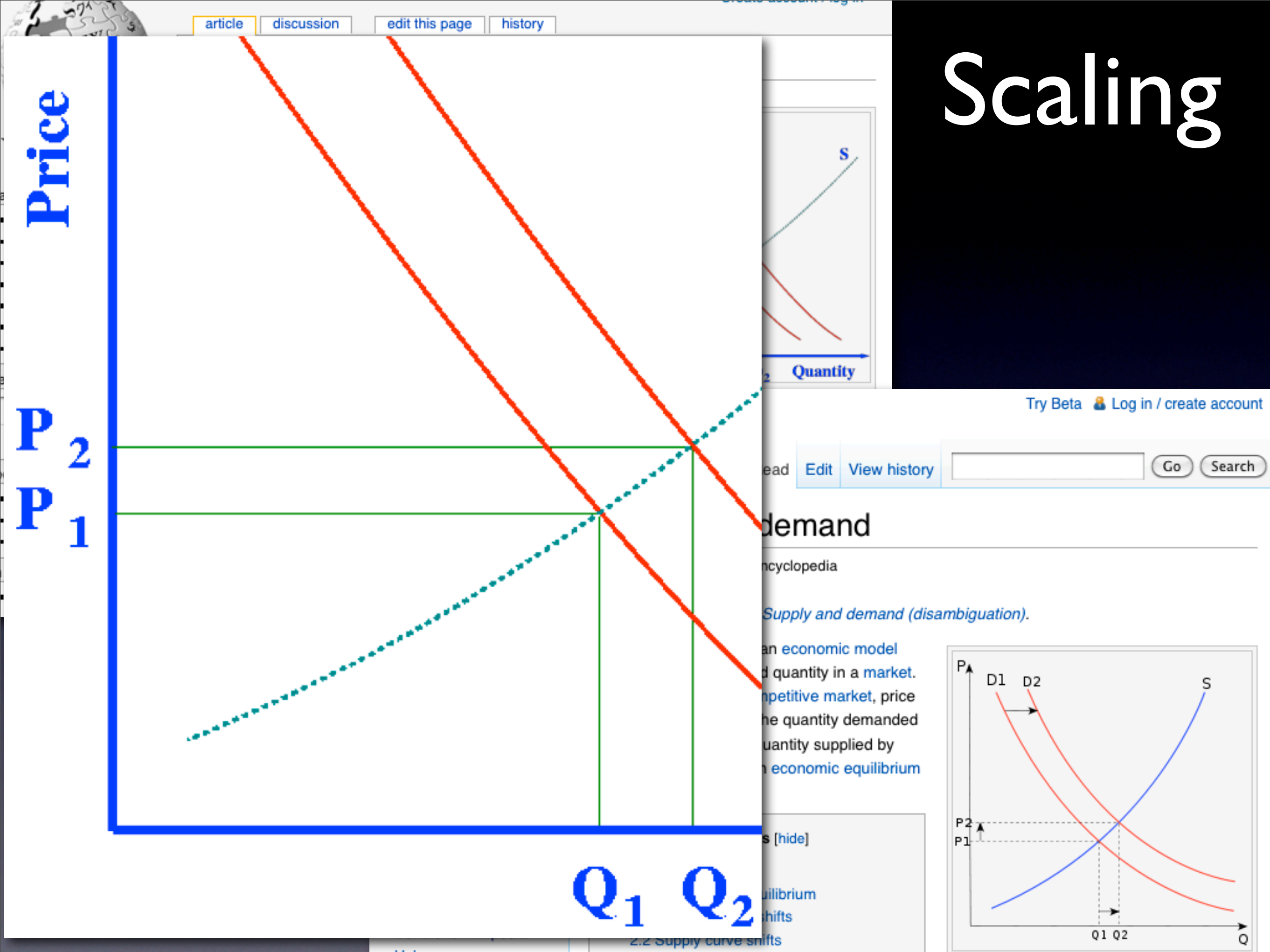
Contents [\[hide\]](#)

- [1 Demand schedule](#)
- [2 Changes in market equilibrium](#)
 - [2.1 Demand curve shifts](#)
 - [2.2 Supply curve shifts](#)

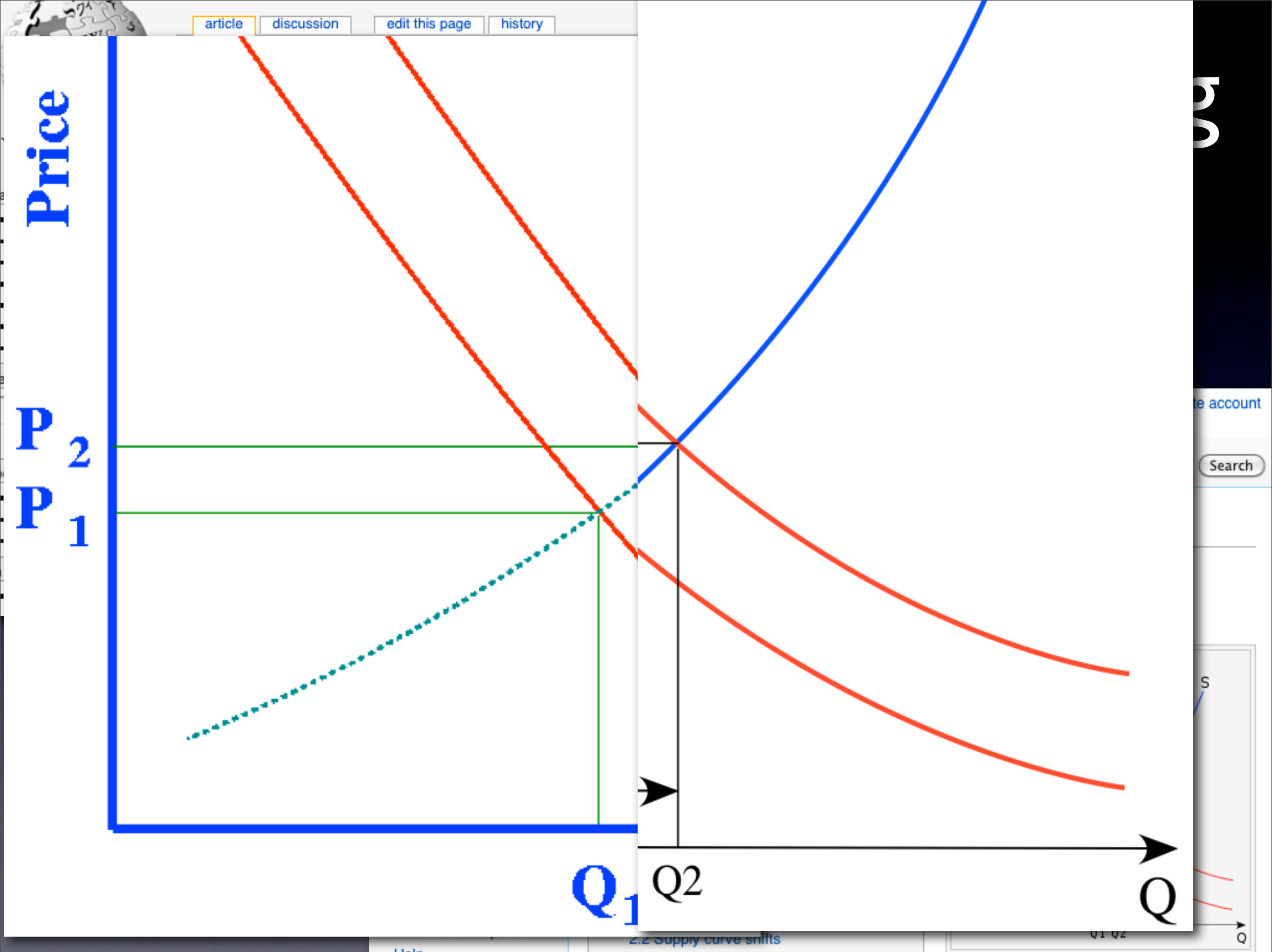


Needs and technology evolve over time. In the early days of Wikipedia many photos and diagrams were created in what were then considered good inline sizes. We now want to support printing, larger high-resolution displays, and simply zooming in to see things more clearly... thousand of old images had to be recreated from scratch to make things future-proof. For photos this is handled by uploading the biggest, cleanest JPEG available and producing scaled thumbnails server-side; for diagrams, maps and logos, vector images can scale down and up as necessary.

Scaling

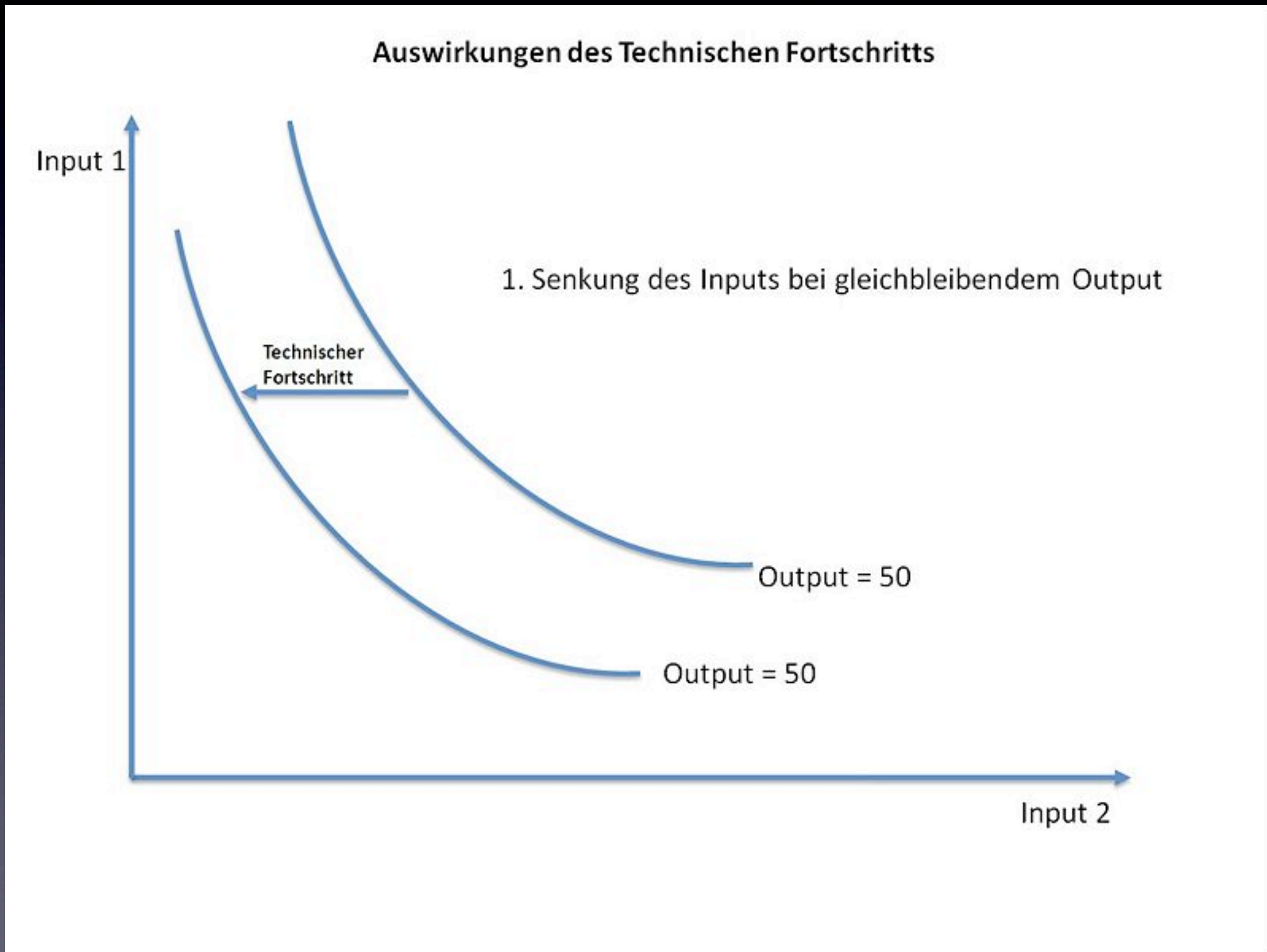


Needs and technology evolve over time. In the early days of Wikipedia many photos and diagrams were created in what were then considered good inline sizes. We now want to support printing, larger high-resolution displays, and simply zooming in to see things more clearly... thousand of old images had to be recreated from scratch to make things future-proof. For photos this is handled by uploading the biggest, cleanest JPEG available and producing scaled thumbnails server-side; for diagrams, maps and logos, vector images can scale down and up as necessary.



Needs and technology evolve over time. In the early days of Wikipedia many photos and diagrams were created in what were then considered good inline sizes. We now want to support printing, larger high-resolution displays, and simply zooming in to see things more clearly... thousand of old images had to be recreated from scratch to make things future-proof. For photos this is handled by uploading the biggest, cleanest JPEG available and producing scaled thumbnails server-side; for diagrams, maps and logos, vector images can scale down and up as necessary.

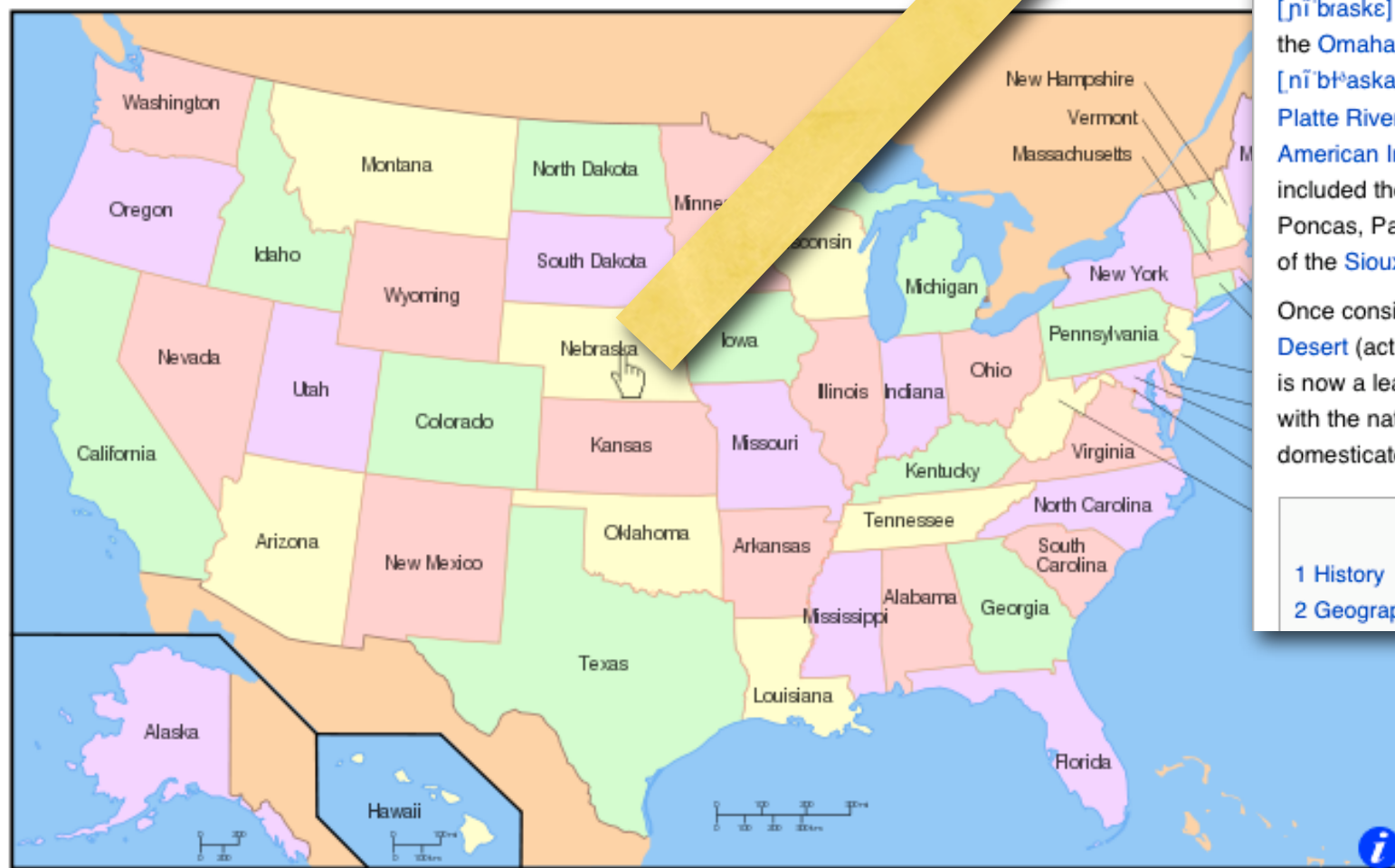
Localization



Wikipedia is a strongly multilingual project, with support for over 250 languages with varying degrees of activity. Maps and diagrams in vector format can be more easily labeled and translated to other languages than flat raster images.

Potential - linking

(workaround OK)



Nebraska

From Wikipedia, the free encyclopedia

Coordinates:  41.5°N 100°W

This article is about the U.S. state. For other uses, see [Nebraska \(disambiguation\)](#).

Nebraska (pronounced /nəˈbræskə/ [ⓘ] [ⓘ] [ⓘ] listen)) is a [state](#) located on the [Great Plains](#) of the [Midwestern United States](#). The state's capital is [Lincoln](#) and its largest city is [Omaha](#).

Nebraska probably gets its name from the archaic [Otoe](#) words *Ñí Brásge*, pronounced [nĩˈbraskɛ] (contemporary Otoe *Ñí Bráhge*), or the [Omaha](#) *Ní Btháska*, pronounced [nĩˈbʰaska], meaning "flat water," after the [Platte River](#) that flows through the state.^[3] [American Indian](#) tribes in Nebraska have included the [Iowas](#), [Omahas](#), [Missourias](#), [Poncas](#), [Pawnees](#), [Otoes](#), and various branches of the [Sioux](#).

Once considered part of the [Great American Desert](#) (actually highly [biodiverse prairie](#) land), it is now a leading [farming](#) and [ranching](#) state, with the nation's largest population of domesticated [llamas](#).

Contents [\[hide\]](#)

- [1 History](#)
- [2 Geography](#)

State of Nebraska



Flag



Seal

Nickname(s): [Cornhusker State](#)
Motto(s): [Equality Before the Law](#)
before statehood, known as
the Nebraska Territory



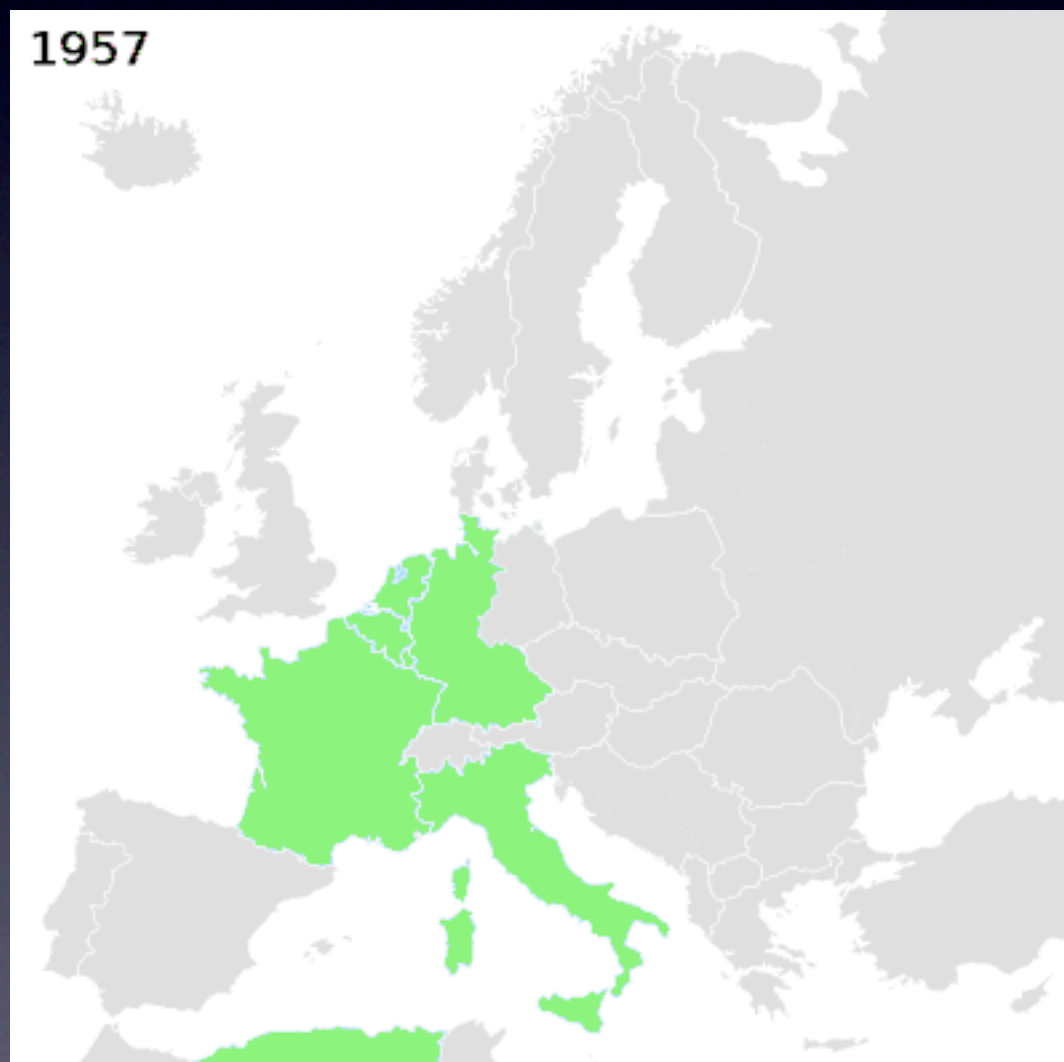
Official language(s)	English
Demonym	Nebraskan
Capital	Lincoln

The ability to embed links directly into SVG images is potentially very attractive for maps, labeled diagrams, etc. For now this can be provided by client-side image maps, but links and hotspots are maintained separately from the source image, complicating authoring.

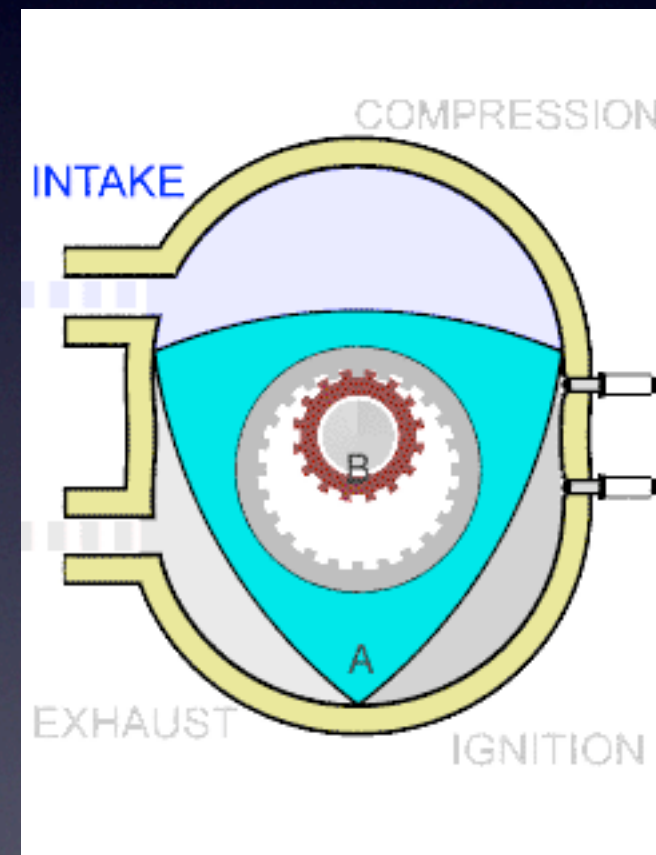
Potential - Animation

Seriously - GIF89?

(workaround awful)



123kb



377kb!

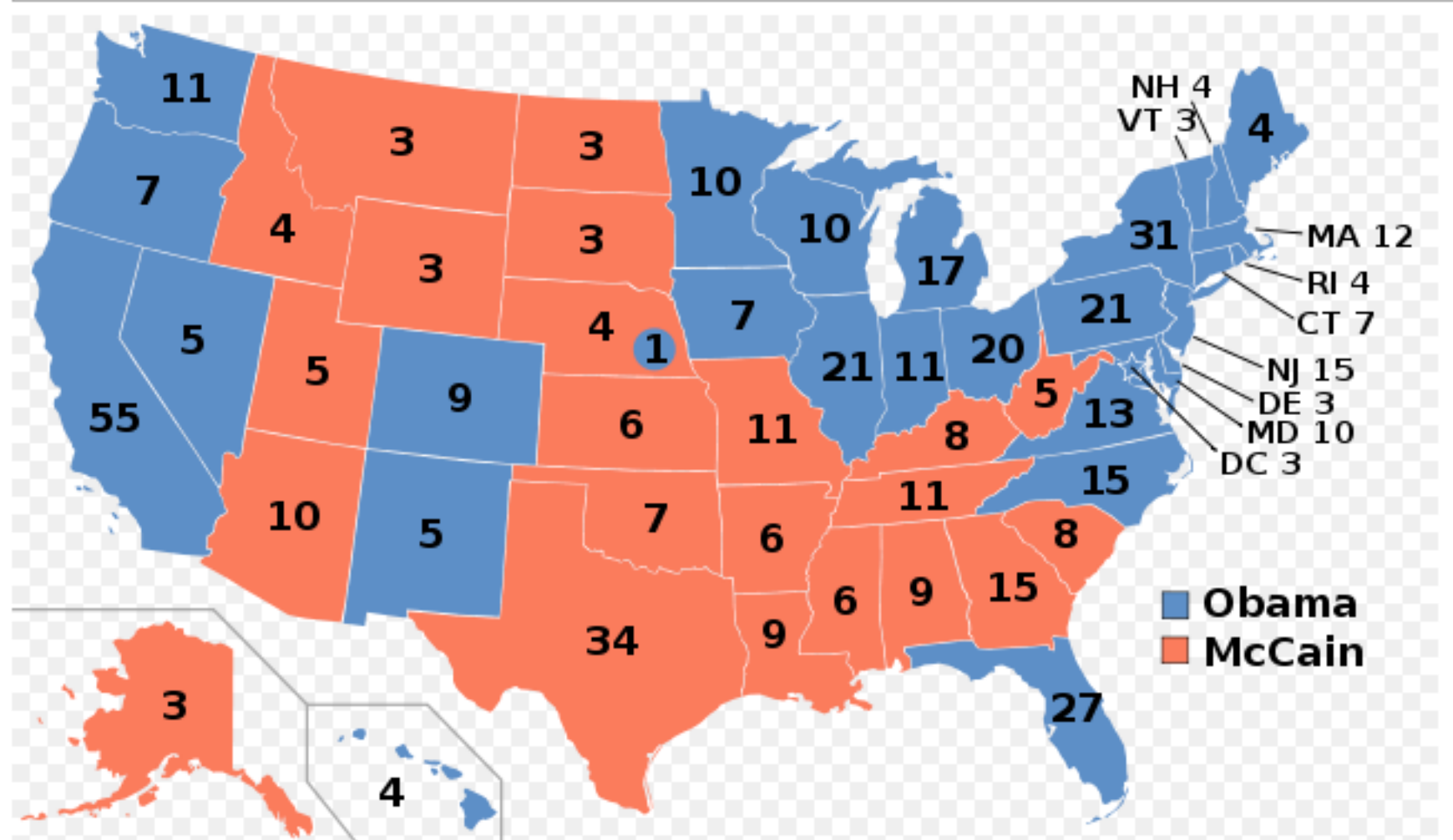
Since rasterization limited us to static files, we put off the issues of animation and interactivity for a future day when support would be better. Unfortunately this has left us with very limited options for animation: Animated GIF and, more recently, video using the free Ogg Theora codec. Neither is well suited for animated diagrams, leading to bloated download sizes or ugly compression artifacts.

Potential - Interactivity

(no workaround)

File:ElectoralCollege2008.svg

[File](#) [File history](#) [File links](#)



US presidential election maps

[1804](#) · [1808](#) · [1812](#) · [1816](#) · [1820](#) · [1828](#) · [1832](#) · [1836](#) · [1840](#) · [1844](#) · [1848](#) · [1852](#) · [1888](#) · [1892](#) · [1896](#) · [1900](#) · [1904](#) · [1908](#) · [1916](#) · [1920](#) · [1924](#) · [1928](#) · [1932](#) · [1936](#) · [1940](#) · [1944](#) · [1948](#) · [1952](#) · [1956](#) · [1960](#) · [1964](#) · [1968](#) · [1972](#) · [1976](#) · [1980](#) · [1984](#) · [1988](#) · [1992](#) · [1996](#) · [2000](#) · [2004](#) · **2008**

We've seen a lot of great demos at this conference showing off interactive charts, graphs, and maps which can be hugely beneficial for educational materials, letting readers get a feel for how something works or slice the data to the part that's most interesting to them. The bad news is that this requires JavaScript; for security reasons we can't let just any potential contributor to expose everyone to arbitrary JavaScript, which could lead to session takeover and compromised accounts. The good news is that it should in most cases be possible to combine crowdsourced static SVGs with a smaller number of JS libraries which have been vetted; common behavior patterns can then be attached to data sources and UI elements provided in the SVG.

Rasterization



Render quality	:)	:	:)
Speed	:(:)	:(
Dependencies	:	:)	:(

The three main toolkits we examined for our server-size rasterization were Batik, GNOME's librsvg, and Inkscape in batch mode. Batik's renderer was most mature, but also the slowest, and depends on a Java environment. Inkscape would render the same as most of our editors would see in their authoring environment, but using a giant GUI app for conversion was heavyweight and harder to maintain. librsvg, designed to render SVG icons and previews throughout the GNOME desktop environment, gave pretty decent rendering in most cases while being much faster and easier to deploy on a Linux server environment.

librsvg bugs?

SVGs fail to render silently if they contain an `<image />` element

Fonts are off in rasterized SVG images on wikimedia sites

Wikimedia Statistics use a nonstandard way of including SVG's in pages.

librsvg workaround for black background on Firefox printing

SVG→PNG and baseline-shift

Incorrect scaling of .svg->.png images.

PNG replacement image for .SVG does not show transparency grid correctly

Strange behaviour with SVG figure

Thumbnail rendering of SVGs broken

Incorrect rendering of stretched text in svg -> png conversion

SVG rasterisation on Wikimedia sites (tracking)

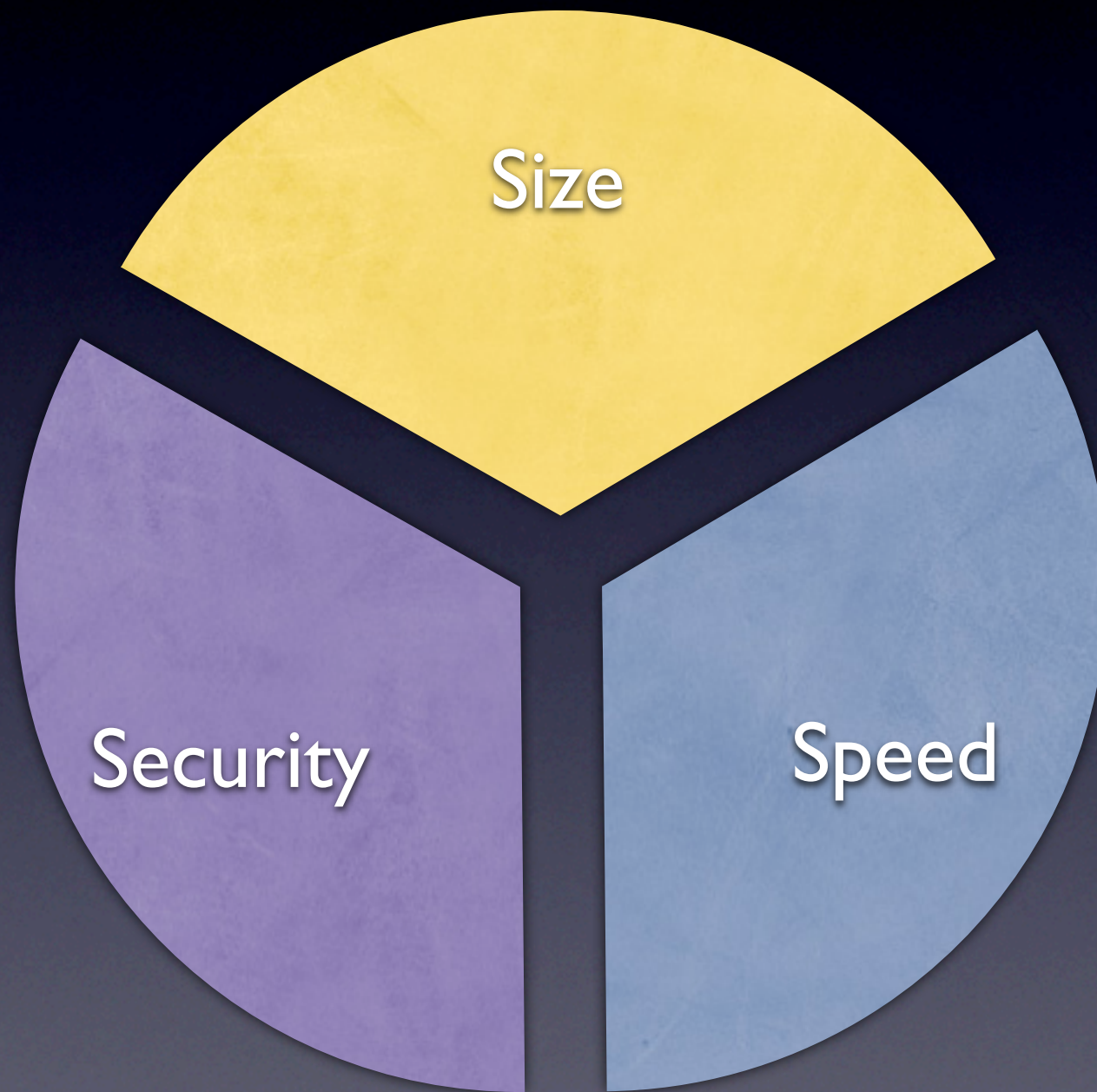
SVG element missing when image is downsized

**ALL YOUR BUG
ARE BELONG
TO ME !**



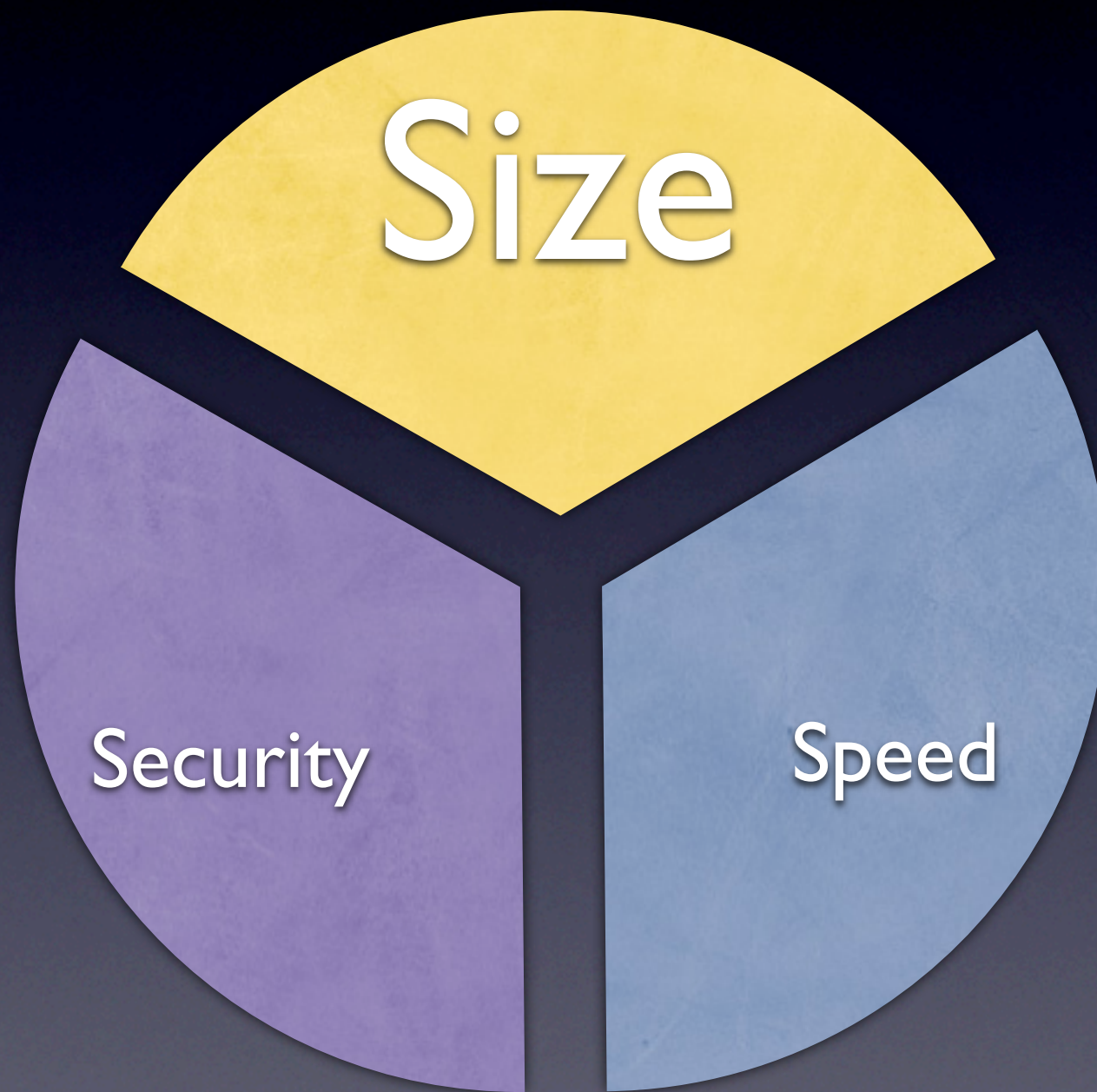
Using a single rendering engine helps keep things consistent, but we've encountered plenty of rendering bugs. Most of these have been tracked down and resolved as librsvg continues to improve, but text layout and font selection and sizing have always been particularly problematic. I'd like to set up an automated test suite running the entire Wikimedia SVG corpus through multiple renderers and calling out significant differences...

SVG challenges



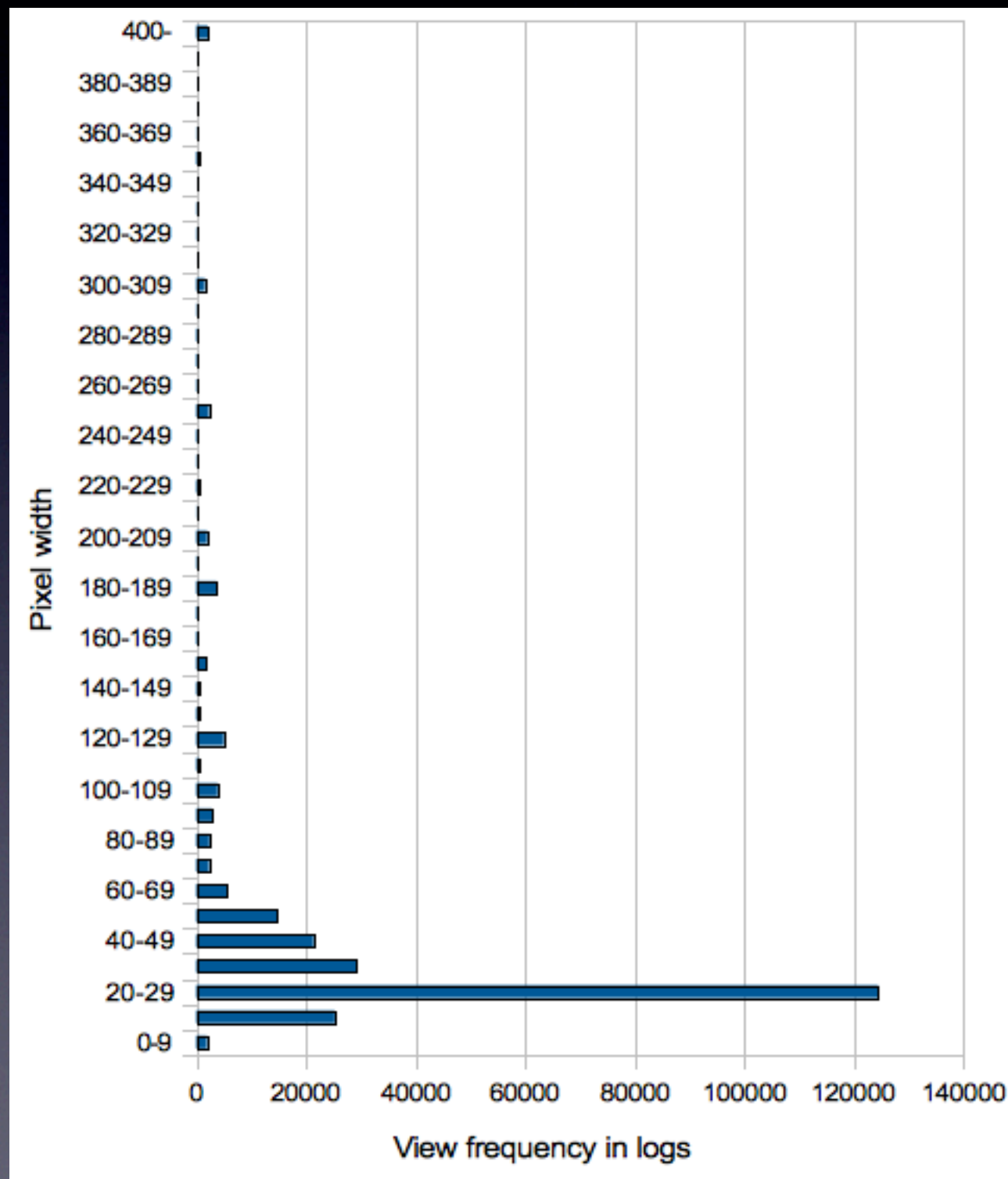
With modern browser support we're finally in a position where we can actually start putting effort into making these new capabilities happen, but there are still some downsides to SVG.

SVG challenges



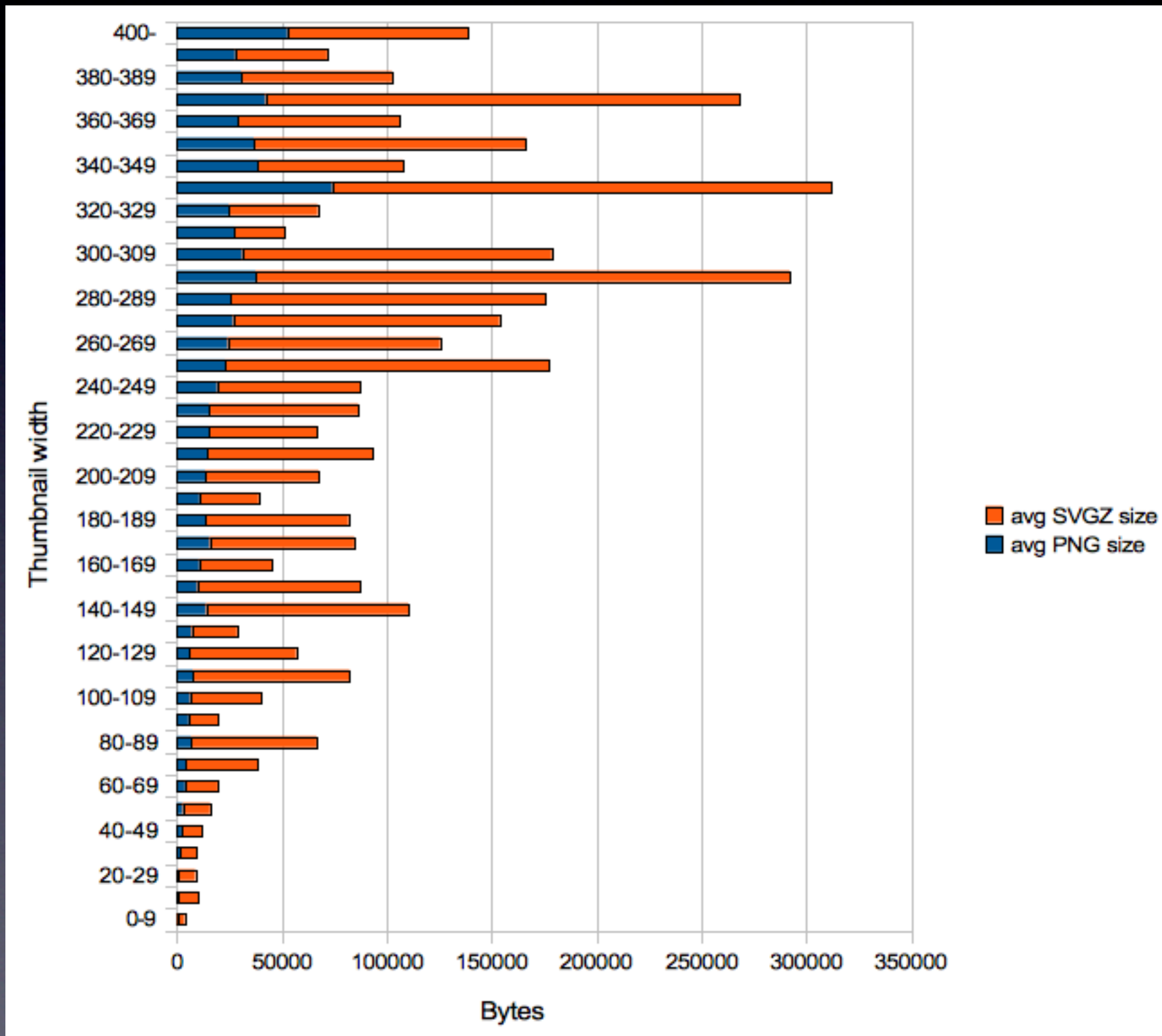
Unlike raster images, SVG source files don't get smaller when you render them smaller. Let's take a quick look at how SVG files are actually used on Wikipedia and how those balance out... I've taken about 12 hours of sampled HTTP log data and pulled the SVG and PNG thumbnails that were actually viewed during this time.

Most popular as icons



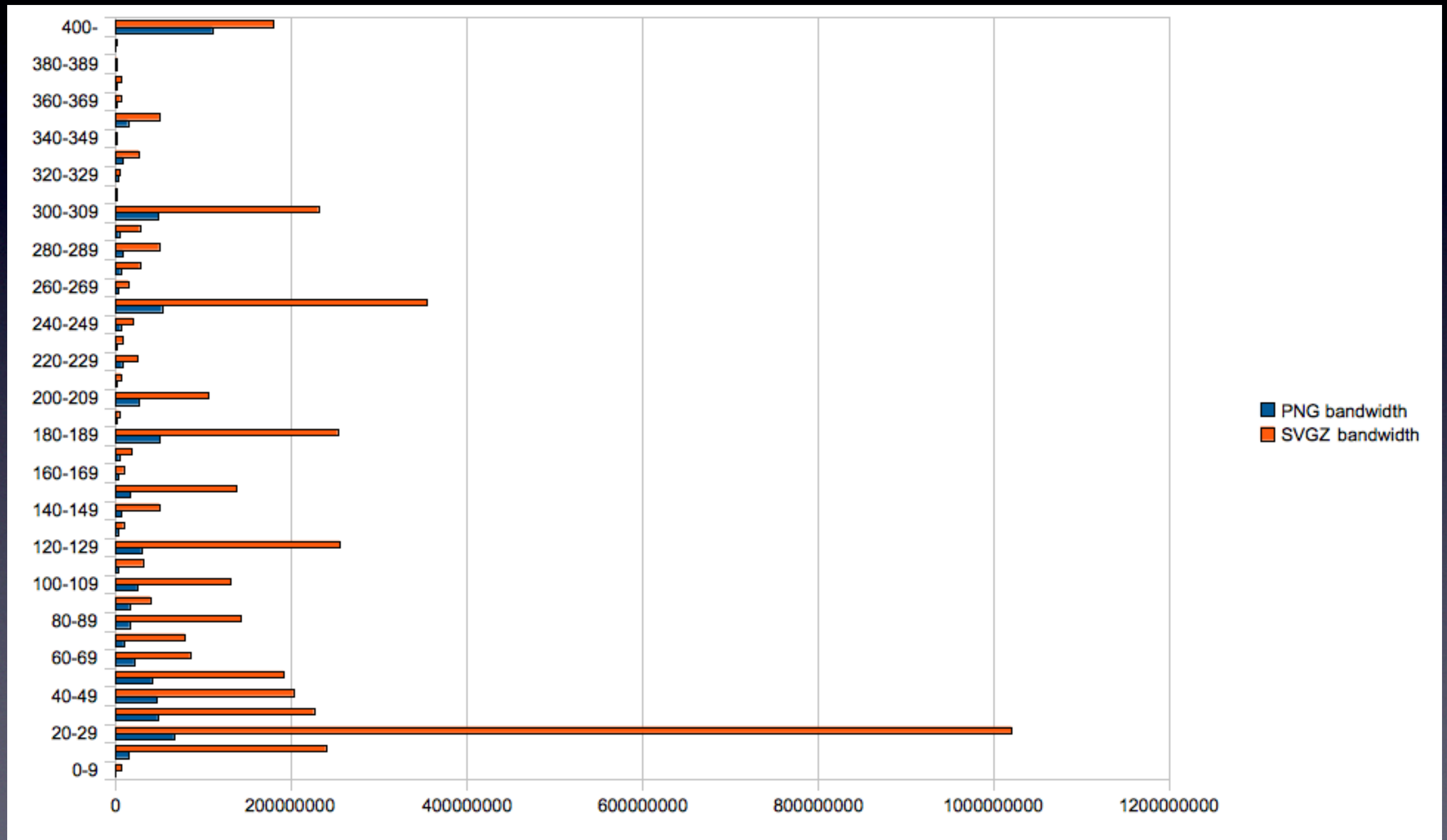
SVG images are most frequently used at small sizes for icons and flags, peaking in the 20–30px sweet spot. Much smaller peaks around 180, 200, 250, and 300 represent a lot of inline maps and diagrams.

Fat SVGs!



Here I'm comparing the byte sizes of the rasterized PNG images we serve now with their source SVG files, broken down the same way by pixel width and weighted by views. In nearly all cases the PNGs are significantly smaller than even the compressed SVG source; right up to the 300px range.

Ow my bandwidth!



If we were to just drop gzipped SVG source files in place of rasterized PNGs, their bandwidth usage would go huuuuugely up.

Too many bytes

				PNG		gzip SVG		
Category group	Avg width	Hits	% hits	Avg PNG	% bw	Avg SVGZ	% bw	SVG bloat
Flags	30px	101,119	39%	570	8%	10,173	24%	18x
Icons	32px	46,108	18%	1,681	11%	7,277	8%	4x
Logos	39px	49,676	19%	2,222	15%	7,511	9%	3x
Maps	192px	7,301	3%	25,599	25%	199,171	34%	8x
All	48px	257,035	100%	2,856	100%	16,669	100%	6x

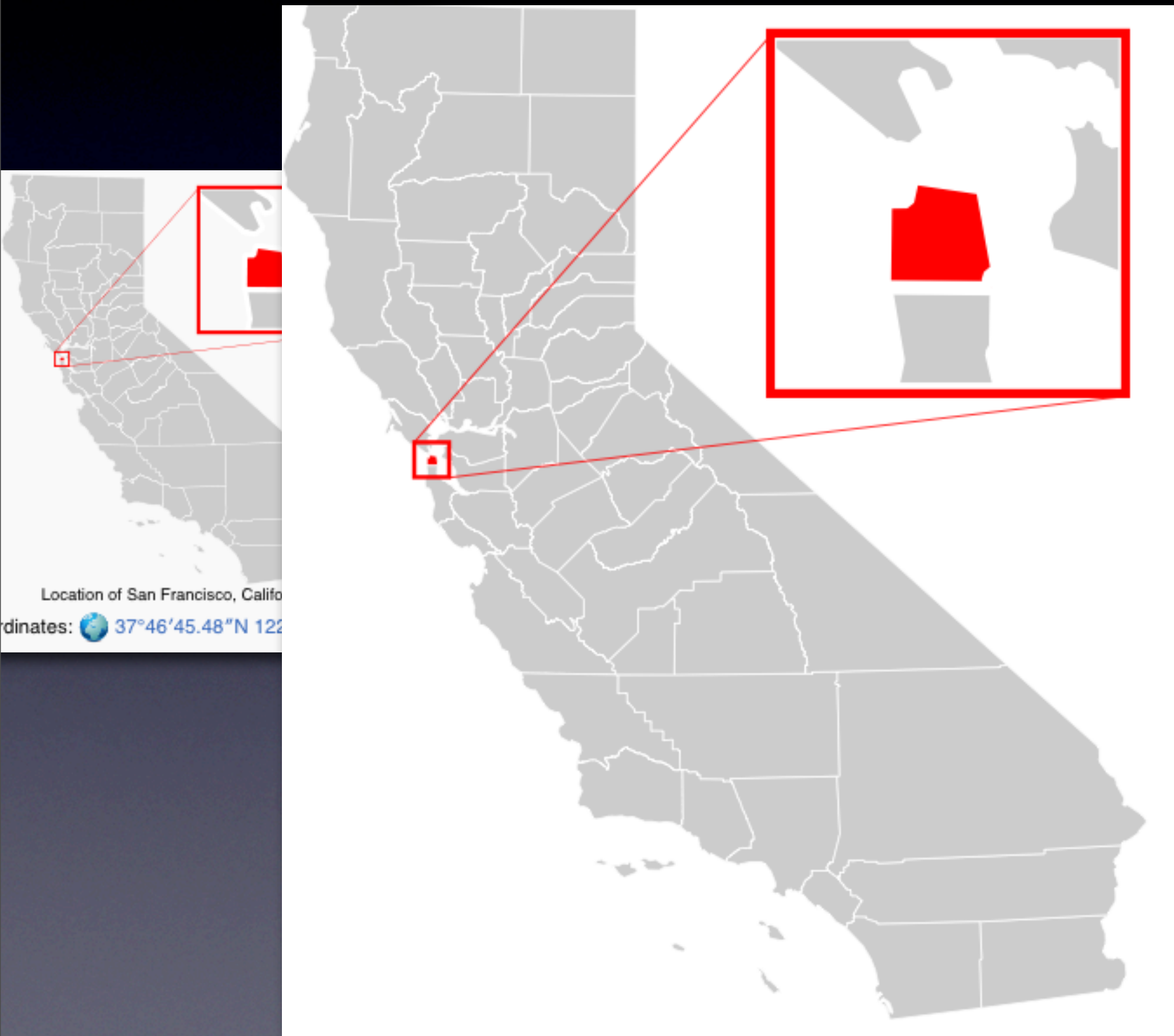
3/4 of SVG views are simple images such as flags, logos, and other icons, usually rendered small. The source is compact -- under 10kb gzipped -- but still several times larger than the rasterized PNGs! Maps are rarer, but disproportionately large... and have even larger source code.

Map madness!



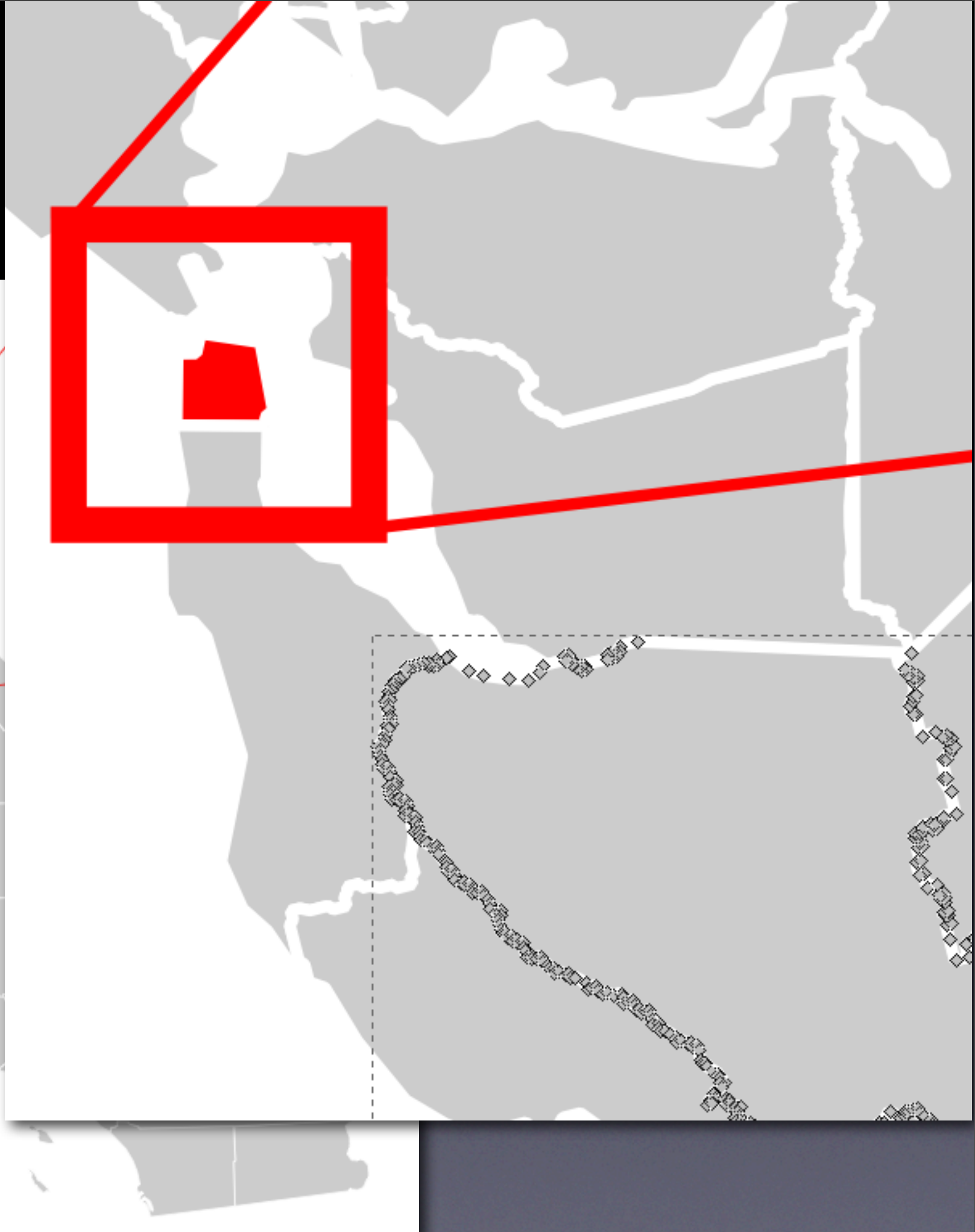
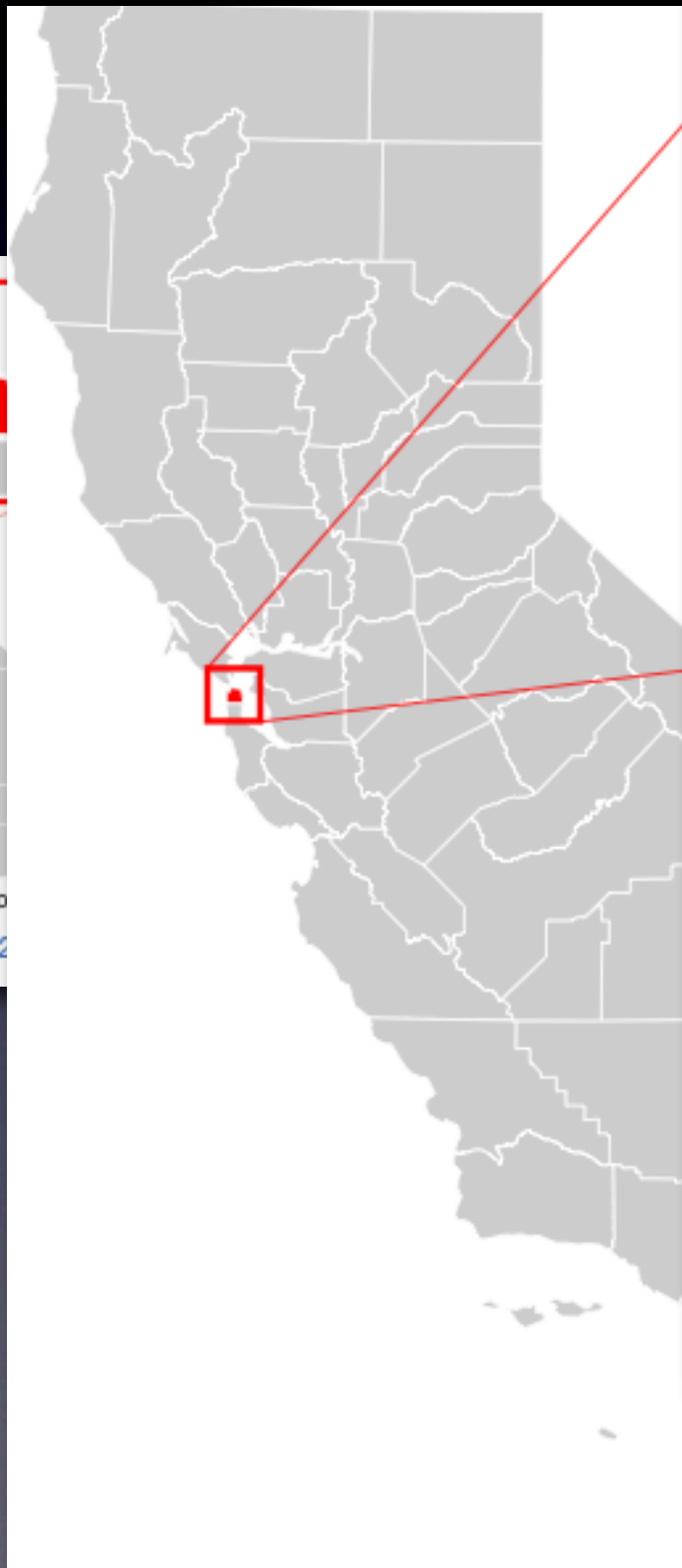
Wikimedians love to err on the side of too much detail! For raster images this is fantastic -- a multi-megabyte 10-megapixel JPEG photo or PNG diagram can be efficiently scaled to a compact thumbnail, and you can easily jump to an insanely-detailed zoomable version. For vector graphics, most of the time that extra detail will be wasted! We only need enough detail to render on screen, including looking good on high-resolution or zoomed screen views and printouts.

Map madness!



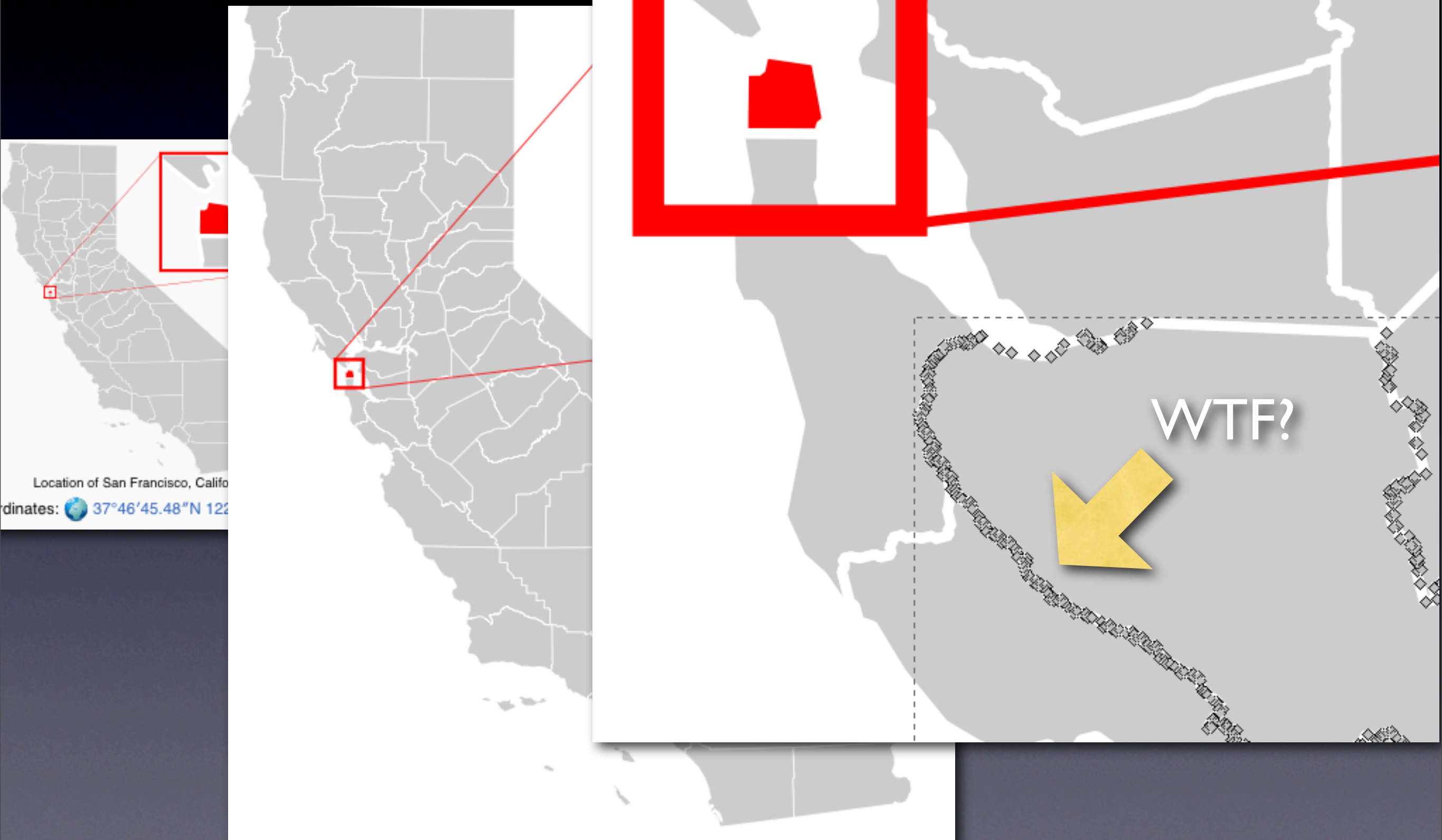
Wikimedians love to err on the side of too much detail! For raster images this is fantastic -- a multi-megabyte 10-megapixel JPEG photo or PNG diagram can be efficiently scaled to a compact thumbnail, and you can easily jump to an insanely-detailed zoomable version. For vector graphics, most of the time that extra detail will be wasted! We only need enough detail to render on screen, including looking good on high-resolution or zoomed screen views and printouts.

Map madness!



Wikimedians love to err on the side of too much detail! For raster images this is fantastic -- a multi-megabyte 10-megapixel JPEG photo or PNG diagram can be efficiently scaled to a compact thumbnail, and you can easily jump to an insanely-detailed zoomable version. For vector graphics, most of the time that extra detail will be wasted! We only need enough detail to render on screen, including looking good on high-resolution or zoomed screen views and printouts.

Map madness!



Wikimedians love to err on the side of too much detail! For raster images this is fantastic -- a multi-megabyte 10-megapixel JPEG photo or PNG diagram can be efficiently scaled to a compact thumbnail, and you can easily jump to an insanely-detailed zoomable version. For vector graphics, most of the time that extra detail will be wasted! We only need enough detail to render on screen, including looking good on high-resolution or zoomed screen views and printouts.

Cleanup squad!



200×236 PNG: **24.5k**



source SVG: **506k** **20x** bigger!

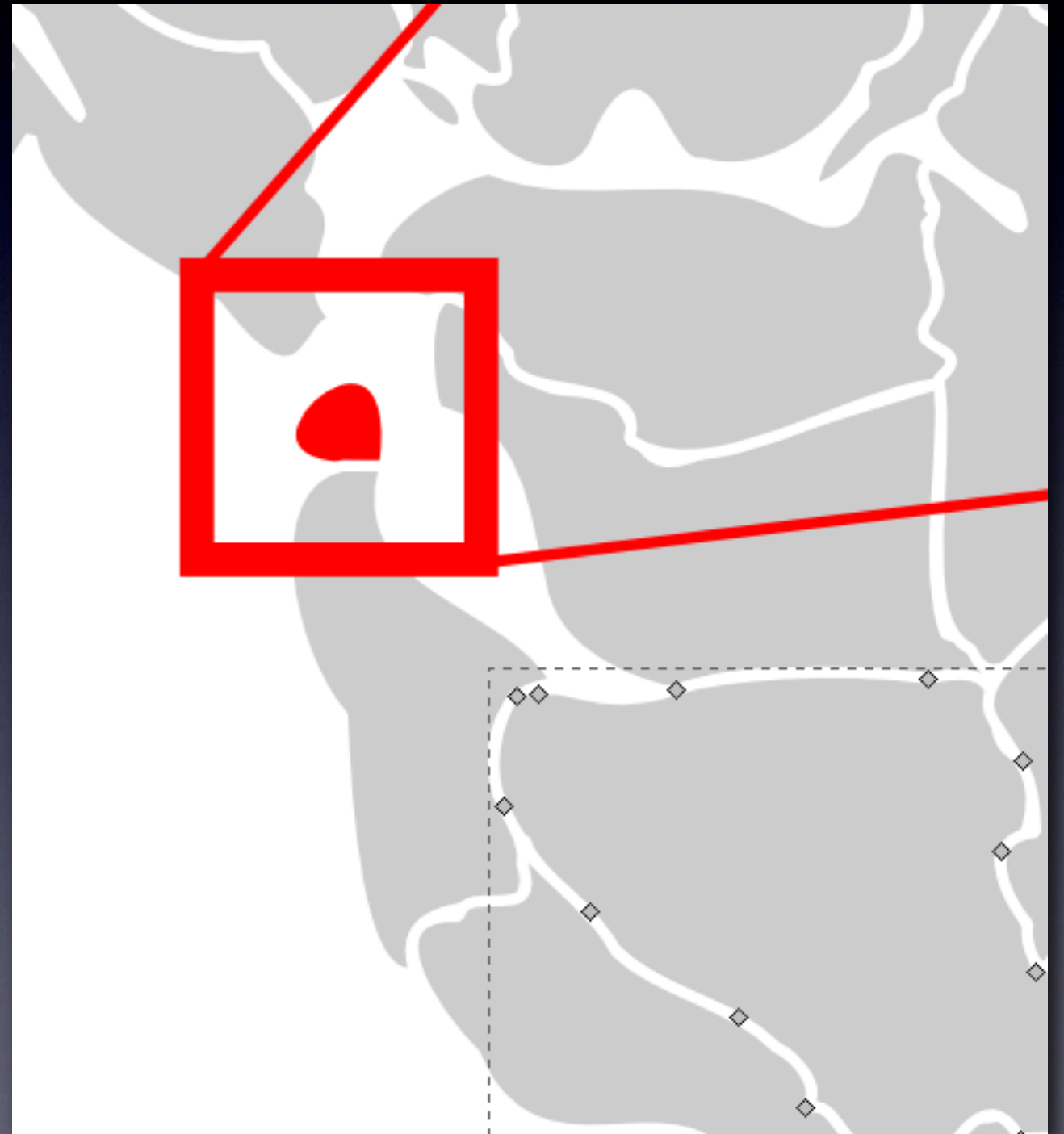
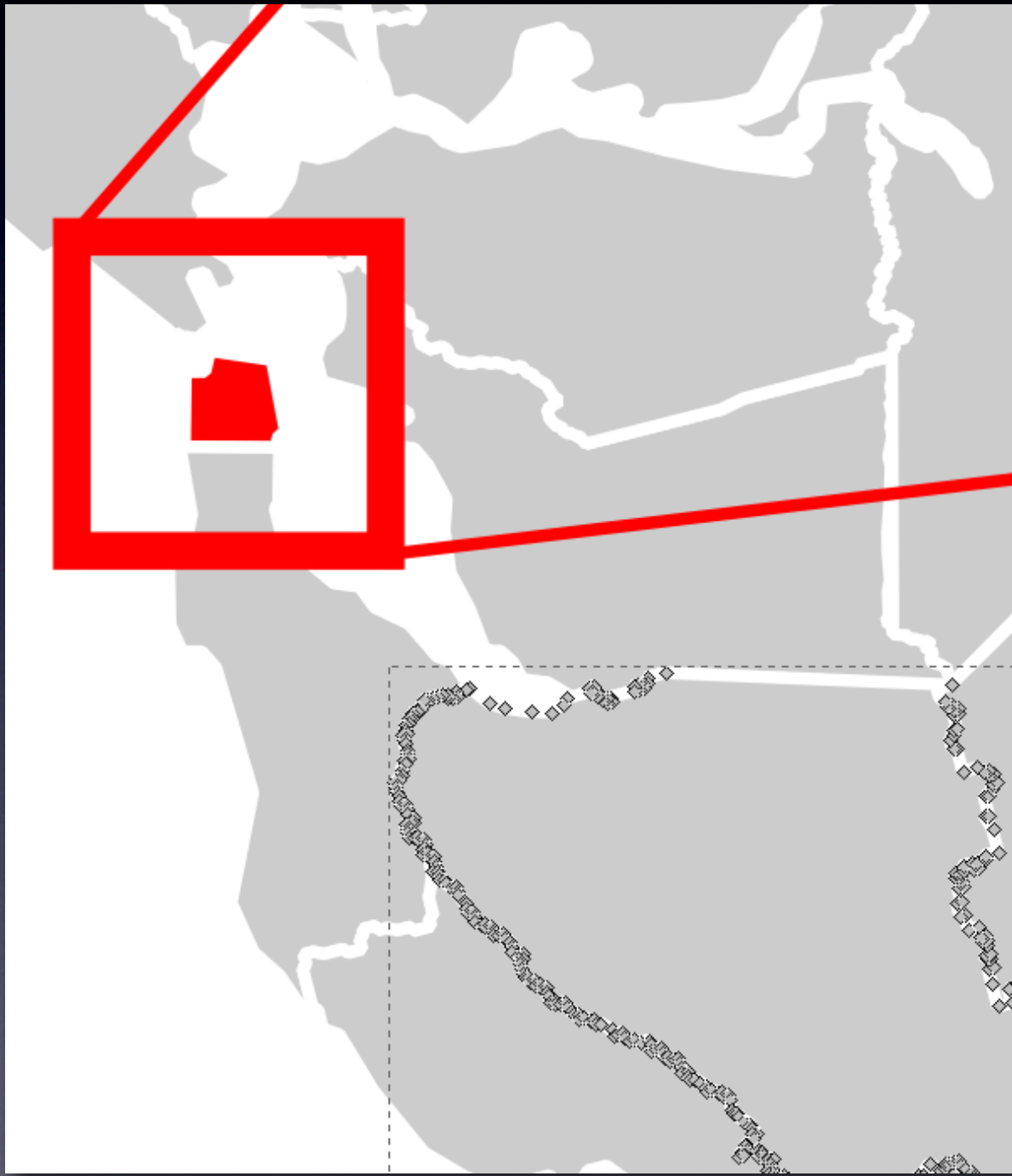
gzip SVG: **239k** 10x bigger

scoured: **237k** 10x bigger

gzip scoured: **49k** **2x** bigger

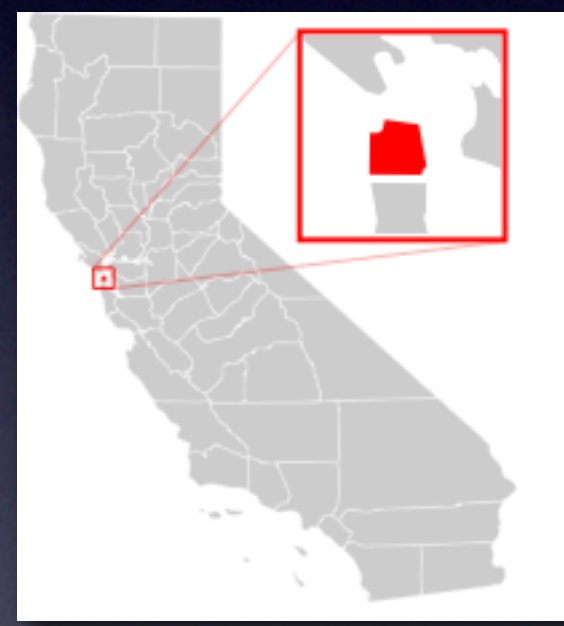
This map appears on the “San Francisco” article at a mere 200 pixels wide; the automatically rasterized PNG weighs in at about 24 kilobytes... but the raw SVG source is a whopping 20 times bigger! Even gzipped we’re 10x the size of the PNG; combining an automated run of Scour and gzip gets us down to 49k with no visible loss. We’re in the right ballpark but still more detailed than we need.

Simplify!



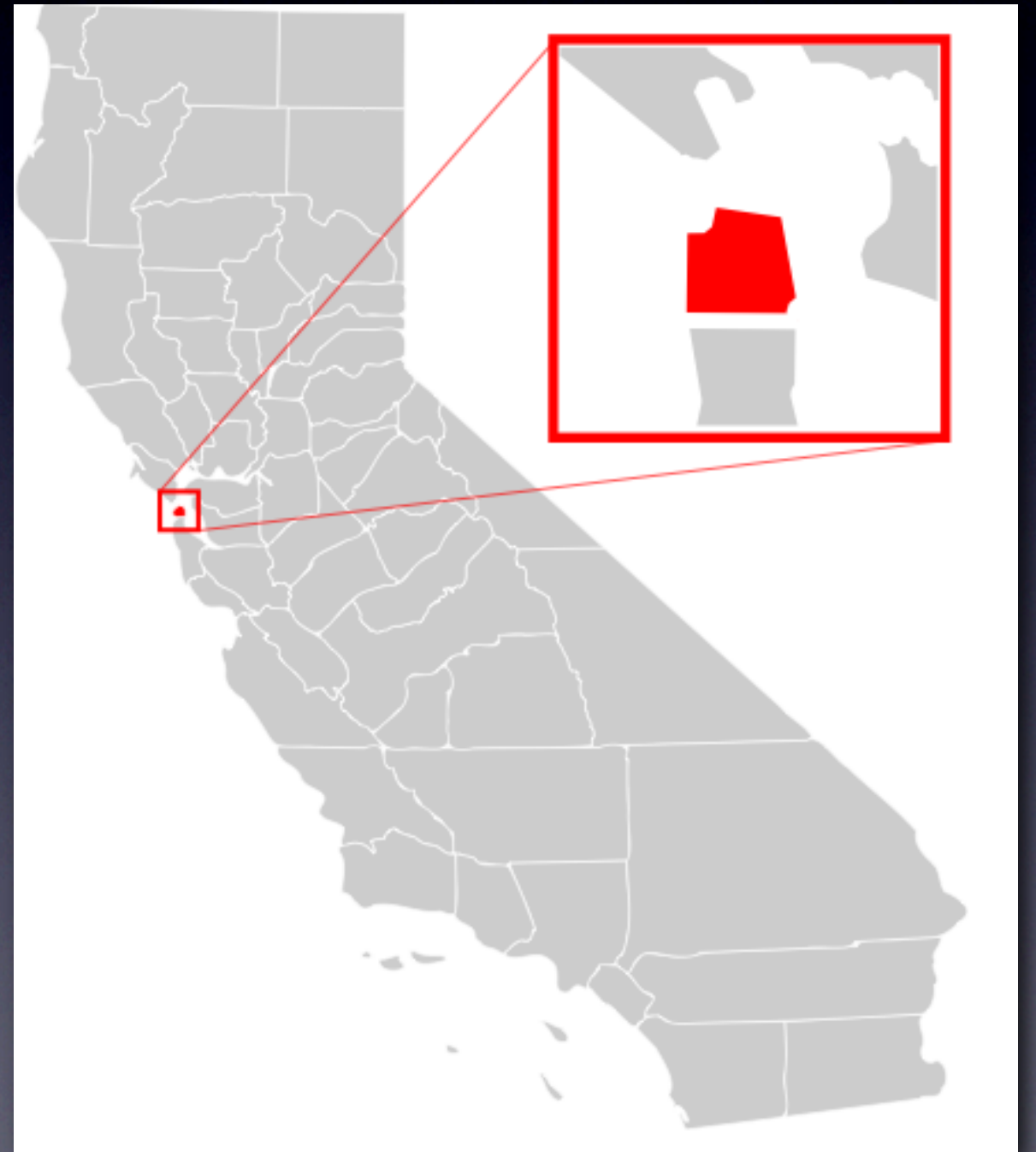
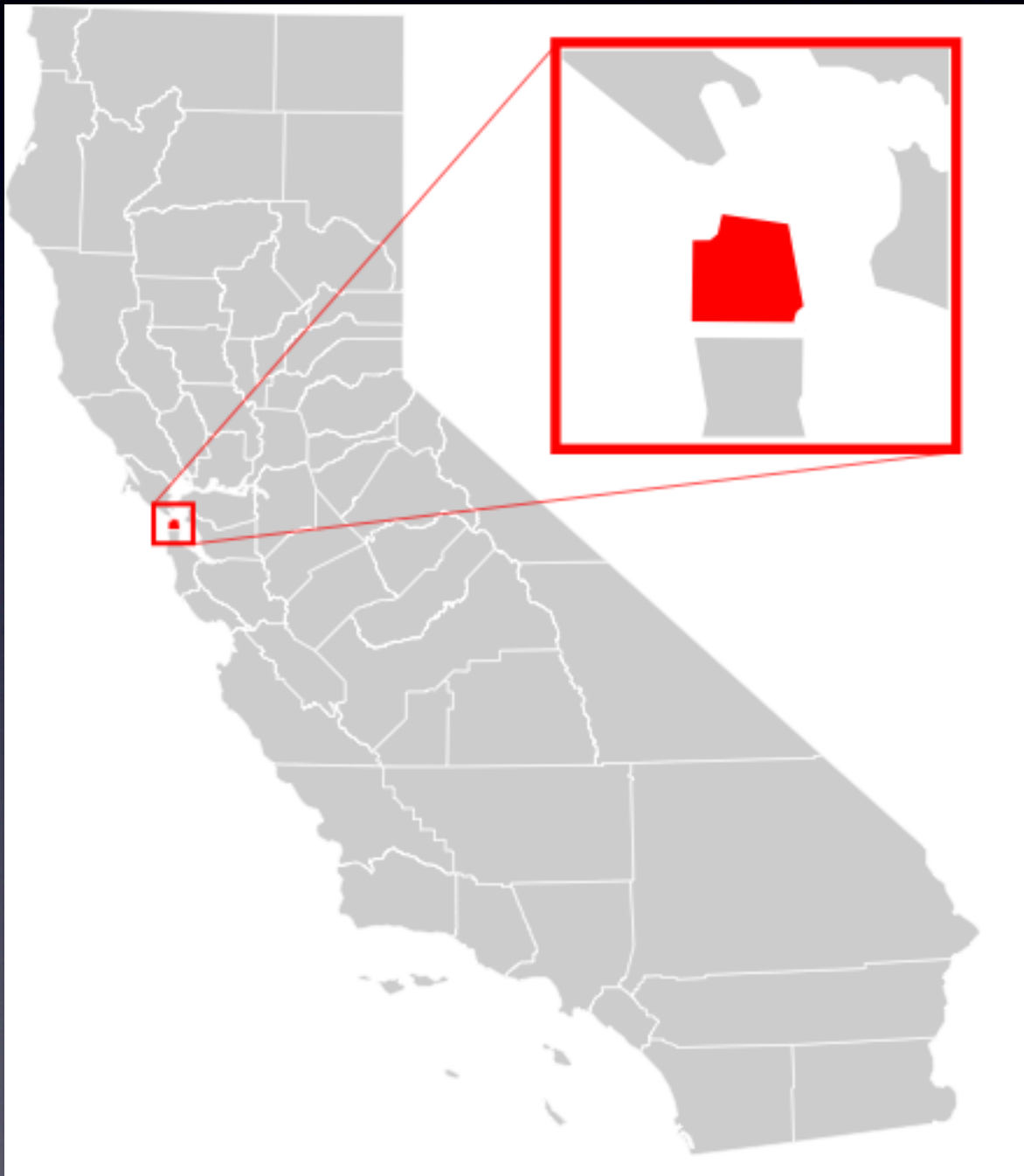
Here we really need a more aggressive approach; we just have a lot of points that aren't necessary for the image as we're actually using it. Scour doesn't yet support path simplification, but Inkscape does have a command for this. When heavily zoomed in we can definitely tell the difference...

Yay!



... but at intended screen sizes, or slightly higher resolutions such as for printing or high-res screens, it looks just fine.

Yay!



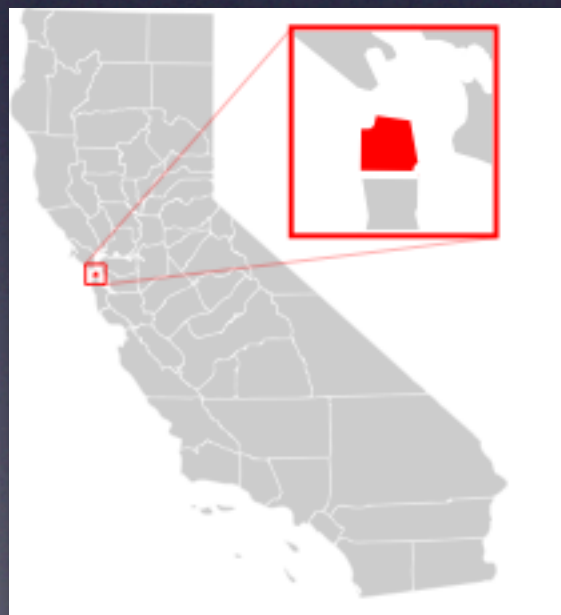
... but at intended screen sizes, or slightly higher resolutions such as for printing or high-res screens, it looks just fine.

Cleanup squad!

simplified



200×236 PNG: **24.5k**



simplified SVG: **295k** **12x** bigger!

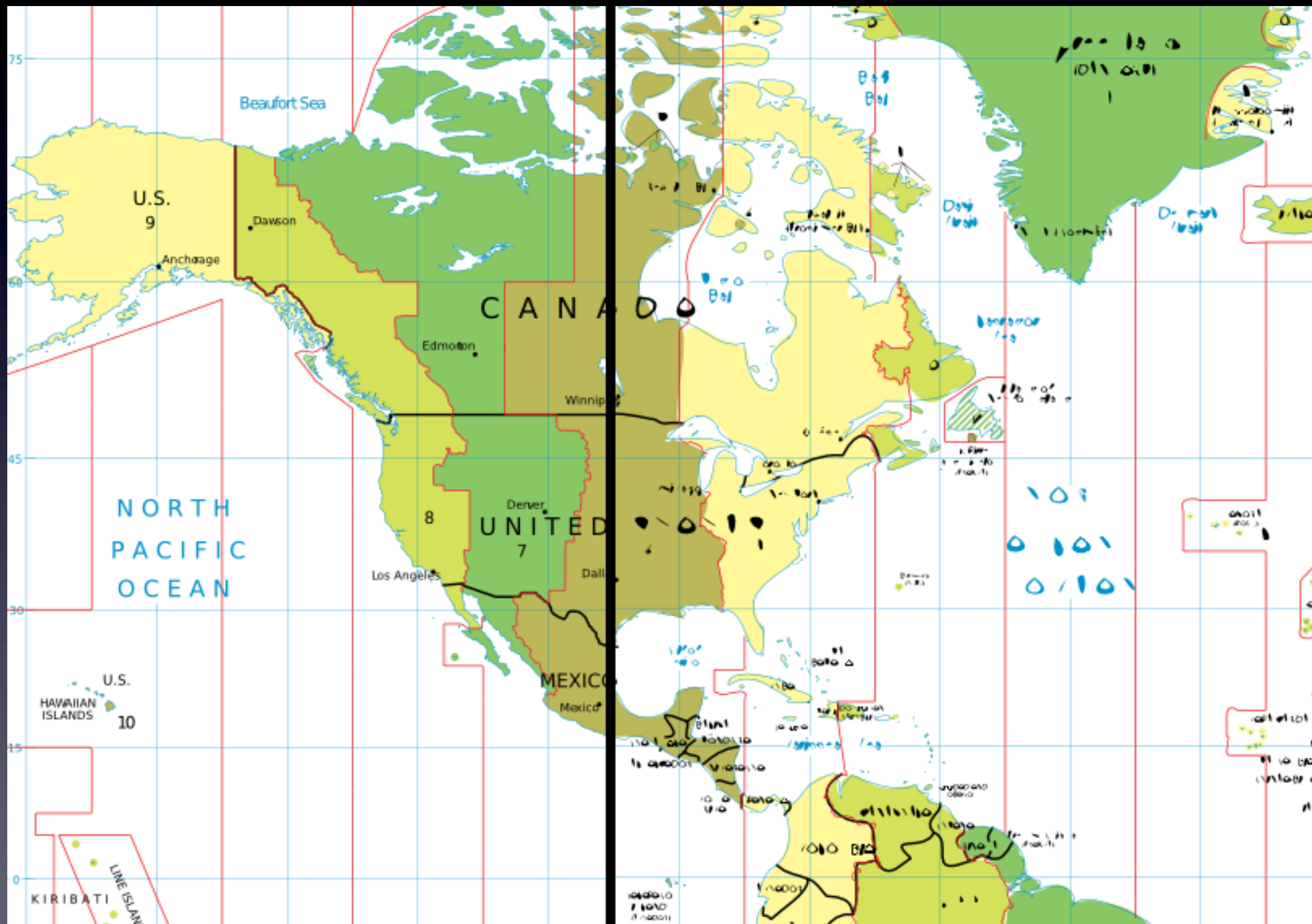
gzip SVG: **209k** 9x bigger

scoured: **42k** 2x bigger

gzip scoured: **19k** **1/5** smaller!

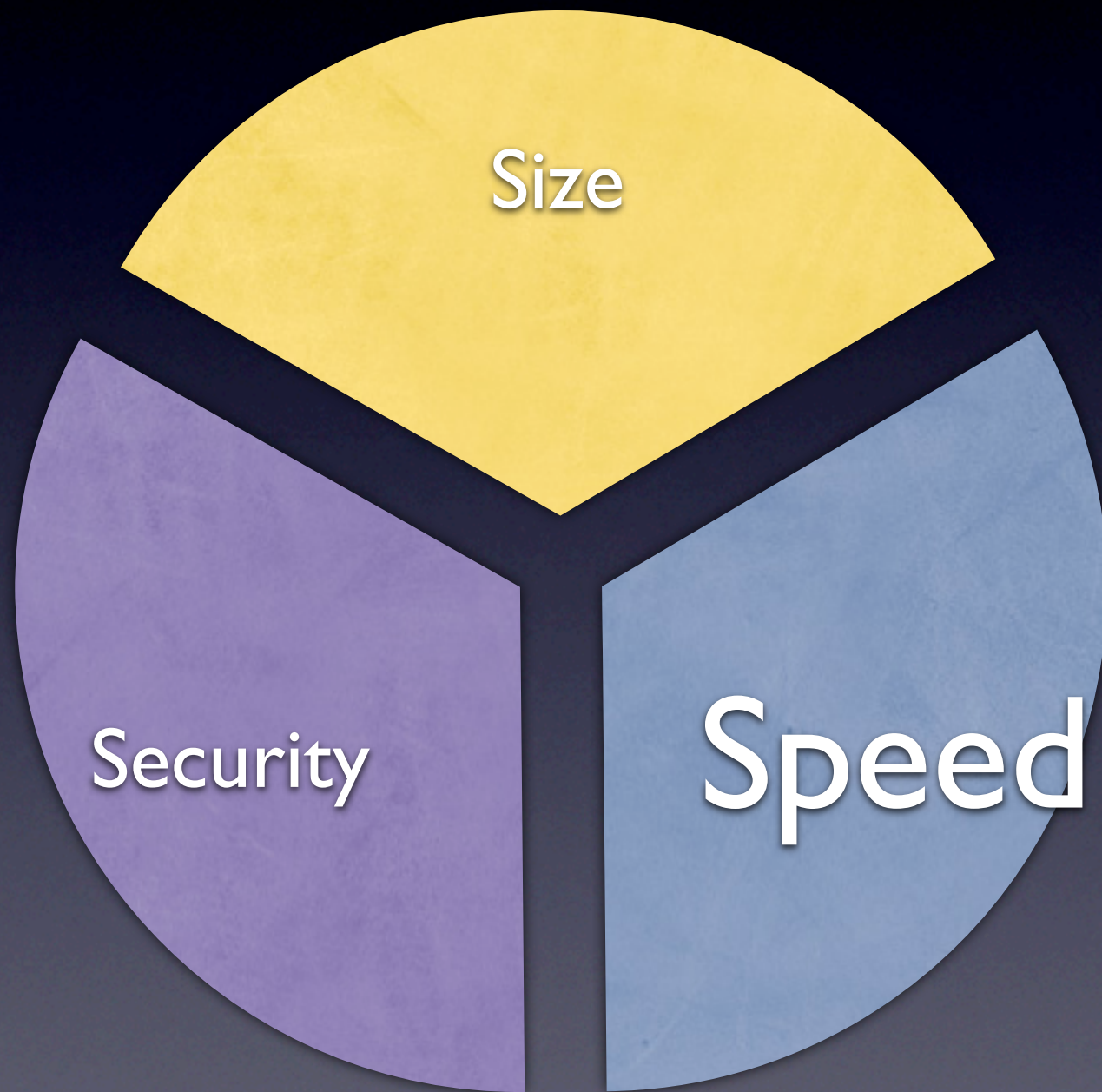
The simplified paths are much more amenable to scouring, and our final gzipped output actually weighed in at 19k, smaller than the PNG. Awesome!

Beware!



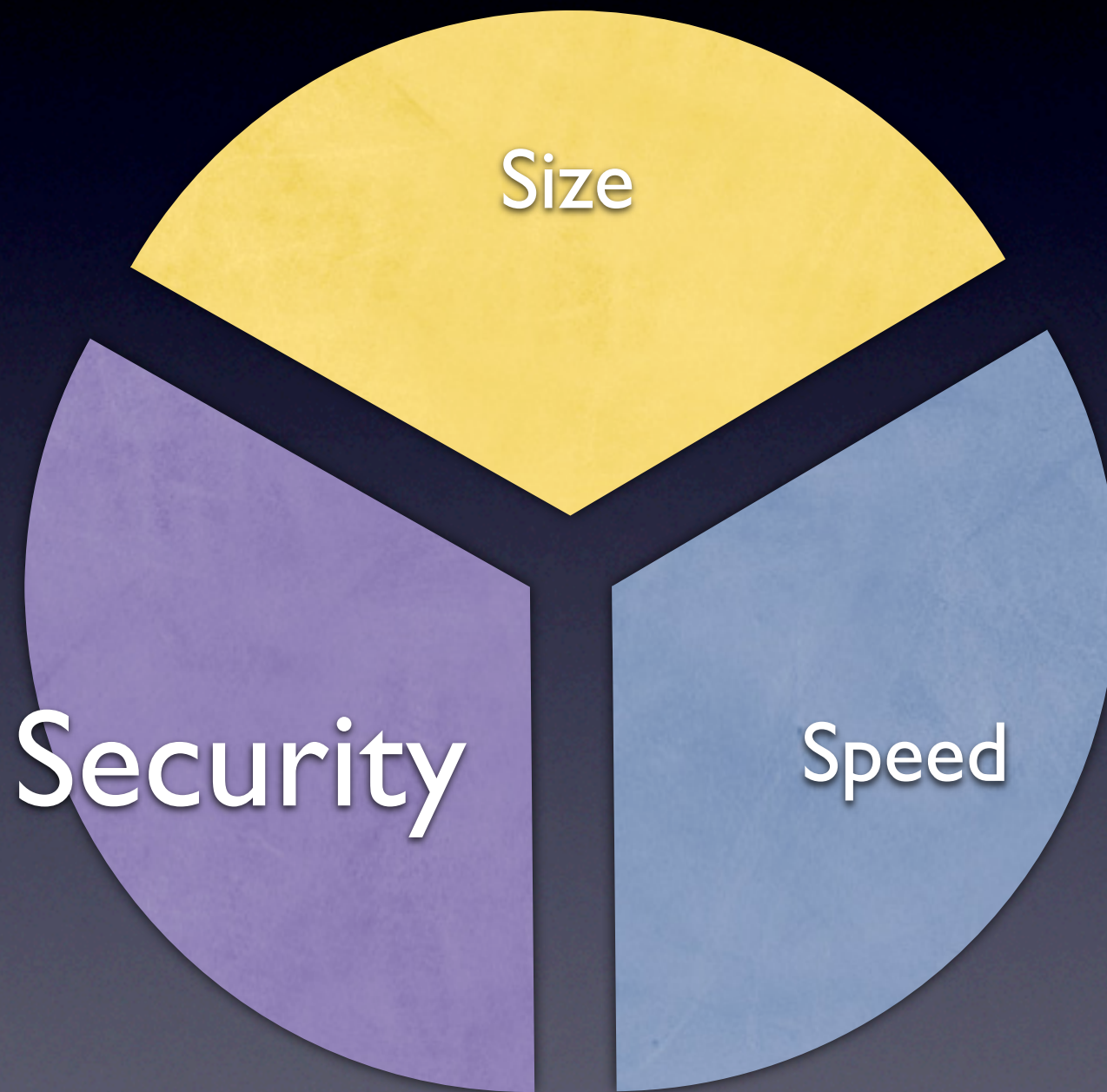
Just applying simplification willy-nilly won't always give good results, though. This 3-megabyte SVG conversion of a detailed PDF timezone map includes fully vectorized text, which looks very bad at the default simplification tolerances in Inkscape! If we do apply path simplification filters post-upload, they need to be placed under user control.

SVG challenges



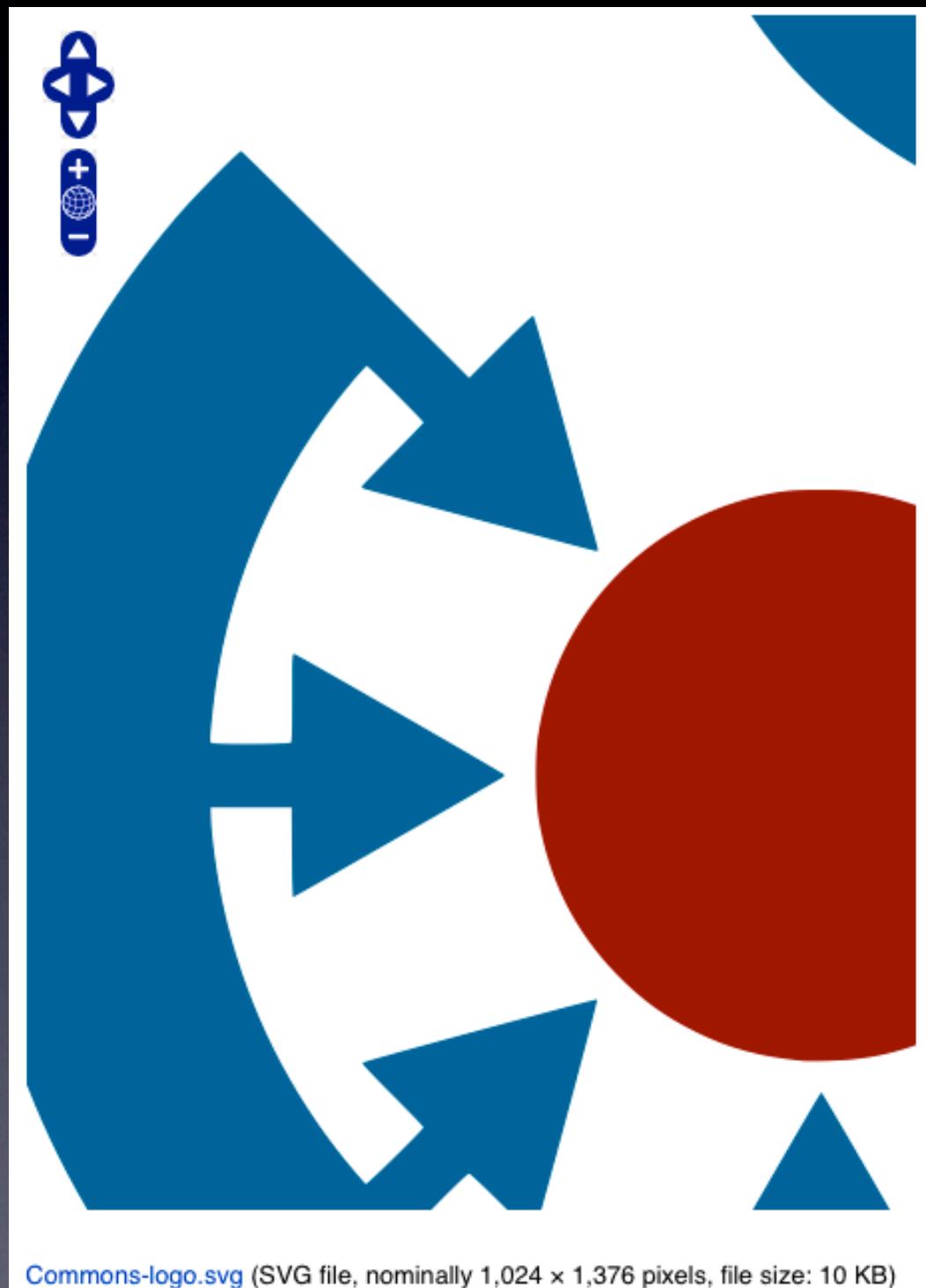
File size affects download speed, but rendering speed is important too. Overdetailed paths slow rendering fairly predictably, but filters can be killer on even simple images.

SVG challenges



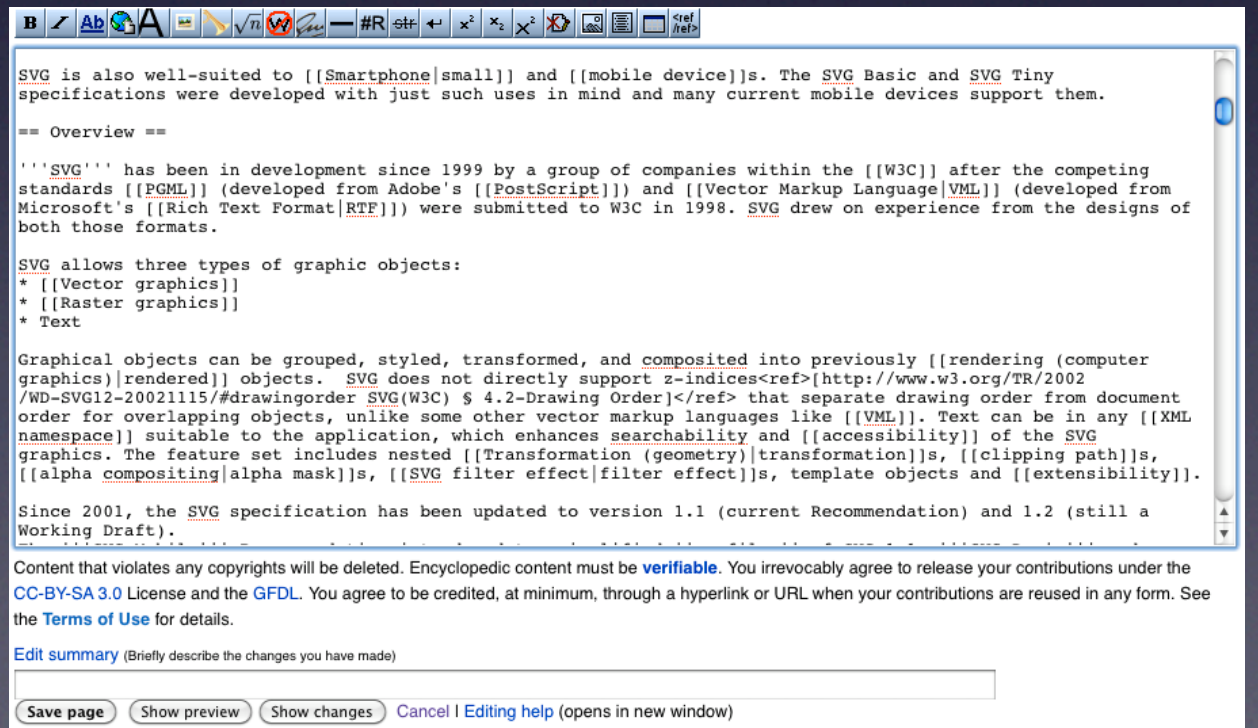
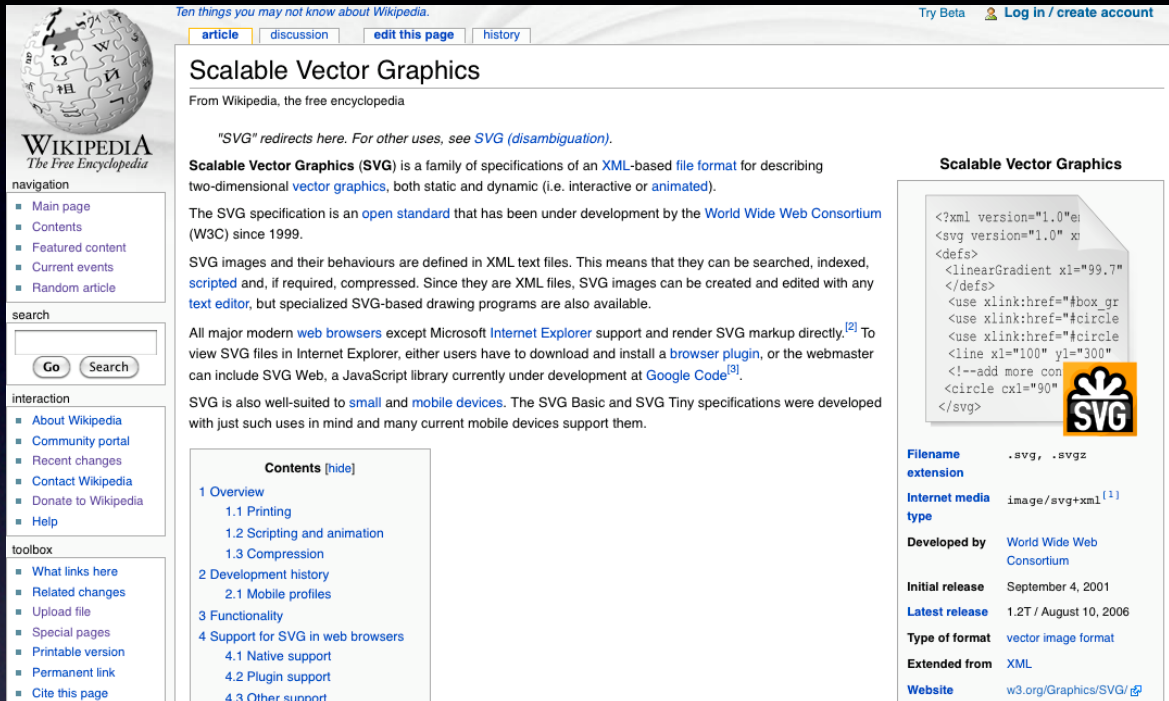
Interactivity requires JavaScript, and animation can benefit greatly from it as well... but access to the browser's JS environment for arbitrarily-uploaded code is a security danger. We already have a model for handling code where trusted administrators can set up global JS code and any user can experiment with their own, which should extend well to attaching libraries providing interactive behavior to "static" SVG graphics.

More hopes?

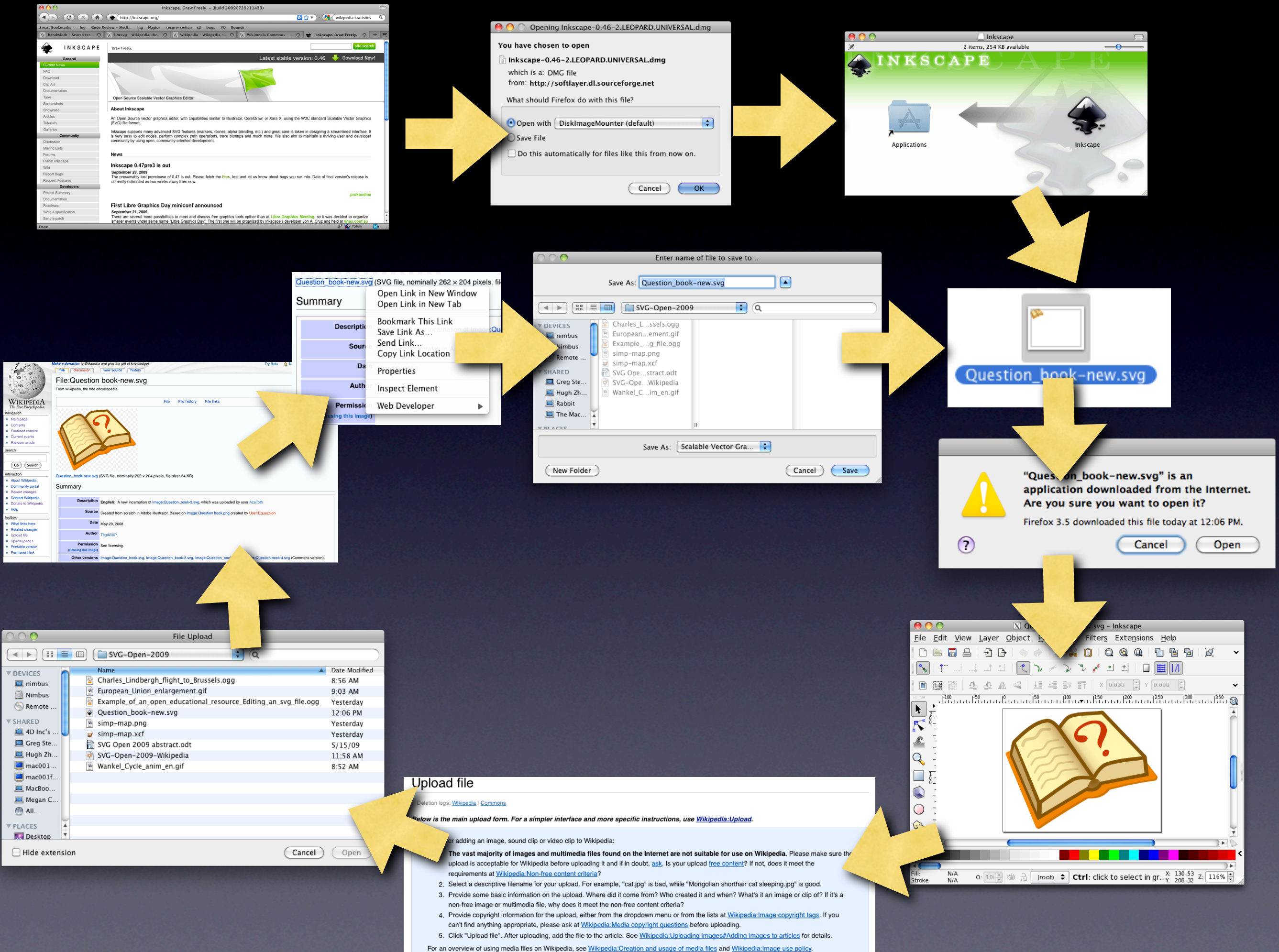


Brad Neuberg showed the other day a demo he whipped up for us to pan and zoom an embedded SVG, including IE support via svgweb. From there we hope to start on interactive tools for localization and charting/graphing, easing some common tasks.

Workflow

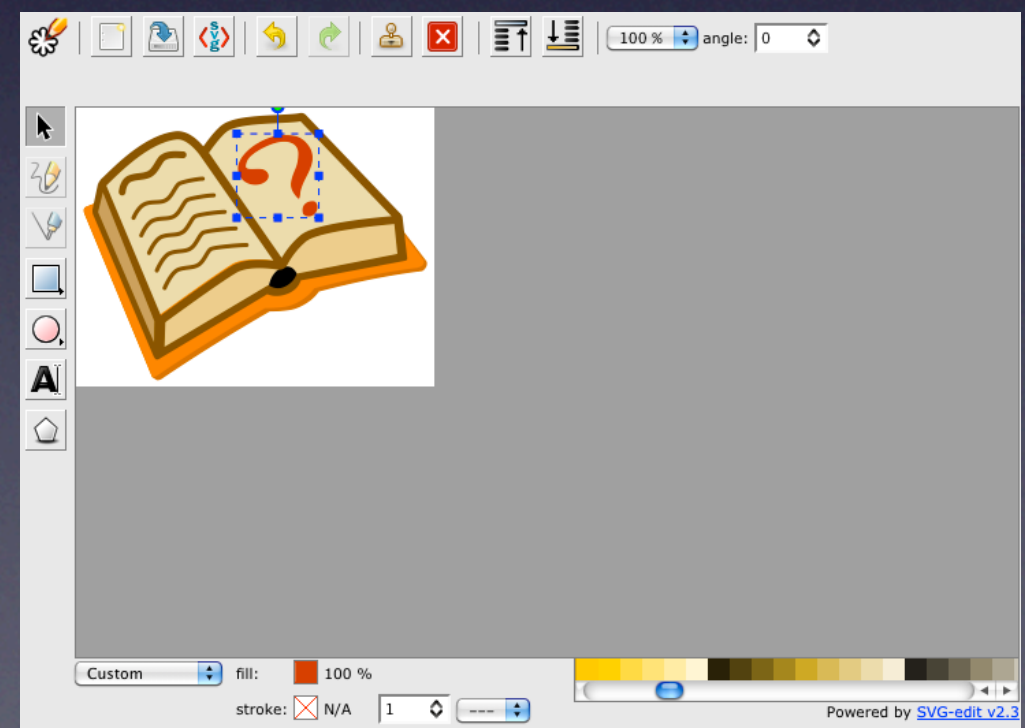
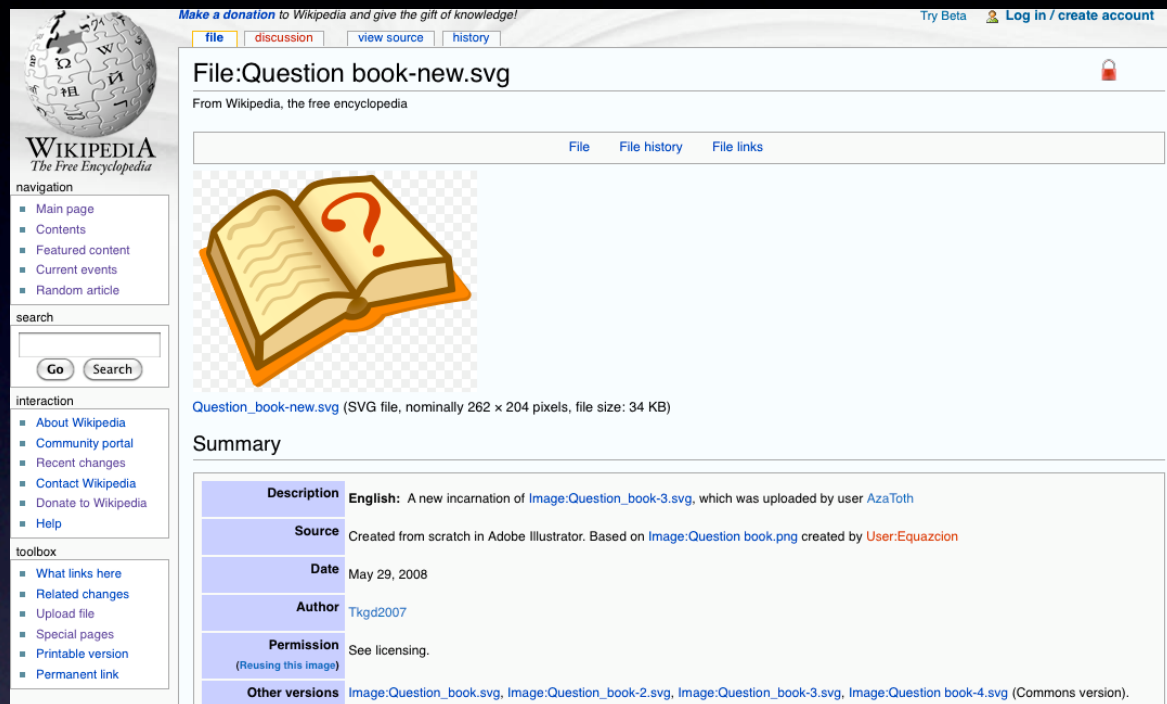


To make participation easy, we try to make editing Wikipedia articles as easy as possible. Text can be edited directly in the browser; with just a couple clicks and some typing you've gone from view to edit and back to view.



Today, the SVG editing workflow looks a little more like this. You have to download and install an editor like Inkscape, then go through a download/open/edit/save/reupload cycle for every change. Ouch!

Workflow?



Ideally, we'd have a decent in-browser SVG editor that can handle most common tasks -- tweaking things here and there, adding labels, and other basic editing. Once SVG-Edit and svgweb get along nicely, it should be pretty straightforward to integrate.

Questions?

http://leuksman.com/pages/presentations#October_2009